JOURNAL OF MATHEMATICAL ANALYSIS AND APPLICATIONS **85**, 409–451 (1982)

A Class of Fuzzy Theories*

ERNEST G. MANES

Department of Mathematics and Statistics, University of Massachusetts, Amherst, Massachusetts 01003

Submitted by L. Zadeh

Contents.

0. INTRODUCTION

At the level of syntax, a flowchart scheme [25, Chap. 4] decomposes into atomic pieces put together by the operations of structured programming [1]. Our definition of "fuzzy theory" is motivated solely by providing the minimal machinery to interpret loop-free schemes in a fuzzy way. Indeed, a fuzzy theory $\mathbf{T} = (T, e, (-)^{\#})$ is defined in Section 1 by the data (A, B, C).

For each set X there is given a new set TX of "distributions on X" or "vague specifications of elements of X." (A)

For each set X there is given a distinguished function $e_X : X \to TX$; "a crisp specification is a special case of a vague one." (B)

For each "fuzzy function" $\alpha : X \to TY$ there is given a distinguished "extension" $\alpha^{\#}: TX \to TY$. (C)

The data are all subject to three axioms. This definition is motivated by the flowchart scheme 1.E. Some fundamental examples are

$$\text{crisp set theory: } TX = X, \tag{D}$$

$$\text{fuzzy set theory: } TX = [0, 1]^X, \tag{E}$$

* The research reported in this paper was supported in part by the National Science Foundation under Grant MCS76-84477.

0022-247X/82/020409–43$02.00/0

probabilistic set theory: TX = set of finite support probability distributions on X, (F)

possibilistic set theory: TX = set of subsets of X. (G)

Space does not permit treatment of program schemes with loops. A number of solutions including "partially-additive theories" are discussed in [5]. Partially-additive theories can also interpret recursive program schemes [6]. Fuzzy set theory, possibilistic set theory, the partial functions theory of 7.5 and many of the "matrix theories" of Section 7 are partially-additive. Crisp set theory is inadequate to deal with loops since an input value may result in an "infinite loop" preventing a crisp outcome.

This paper offers a language to compare theories. For example, (D), (F) are noise-free, (E), (G) are not. Examples (D), (F) have crisp points while (E), (G) do not, but every theory has a largest canonical subtheory with crisp points which for (E) is related to the normalized fuzzy variables of [11] (see 1.15). All four are commutative theories which are antireflexive, faithful, propositionally complete, conditional-complete and which satisfy the eigenstate condition. Our formulation provides a "Boolean logic" for every commutative theory. The fuzzy set complement operation of [28] $[0, 1] \to [0, 1]$: $x \mapsto 1 - x$, is our complement for the theory of (F)!; whereas our complement for the theory of (E) is $[0, 1] \times [0, 1] \to [0, 1] \times [0, 1]$: $(x, y) \mapsto (y, x)$ in agreement with [11].

Much investigation in "non-standard set theory" begins with the premise that a non-standard set is a representation in a non-standard logic of truth values, be it for observables in quantum statics [16, p. 98] or for fuzzy sets [12, 28]. If the models are allowed to vary at all, focus is on the axiomatic structure of the logic which is usually viewed as a lattice. Recently, however, topos theory (see [18]; reviewed in [23]) has demonstrated that the generalization from two-valued logic to the Brouwerian logic of intuitionistic set theory follows from axioms on more primitive structural features (this amounting to no more than a precise statement of "a subset of $X \times Y$ is a function from X to the subsets of Y") thereby deriving the concept of Brouwerian lattice rather than positing "intuitionistic logic" from the outside. In the same spirit, the simpler principles motivating fuzzy theories are powerful enough to represent distributions as truth-valued functions as follows.

Denote by $\mathscr{T}$ the set $T\{\mathbf{true}, \mathbf{false}\}$ of fuzzy truth value. Feeling that the equality of vaguely specified things is also vague, we derive "equality maps" of form $eq_X\colon TX \times TX \to \mathscr{T}$ in terms of which the "degree of membership of $x \in p$" for $x \in X$ $p \in TX$ is defined by $dm_X(x, p) = eq_X(e_X(x), p)$. This gives rise to the above-mentioned representation $p \mapsto dm(-, p)$ of TX in the "proposition space" $\mathscr{T}^X$.

We show in Section 6 that (for commutative theories such as the fundamental examples mentioned) every Boolean polynomial extends canonically to $\mathscr{E}^X$ by virtue of a general "fuzzification principle" so that there is always a logic of propositions. In this framework we extend a result of Eilenberg from the possibilistic theory (G) to arbitrary commutative theories, to obtain the metatheorem that every equation involving the same set of variables in each term without repetitions (such as de Morgan's law $(x \vee y)' = x' \wedge y'$, but not idempotency $x \vee x = x$) must be true in the logic of propositions. (I am aware of Eilenberg's results from talks that he gave circa 1970 on the "linear theories" that are promised in the preface of [9], but there is nothing in print to my knowledge). Even though the representation of TX in the proposition space is injective in most cases, there is no reason why the generalized Boolean operations should map distributions to distributions. Indeed, in the primal motivating example of crisp set theory (D) the proposition space is the usual Boolean algebra of subsets but a Boolean polynomial applied to singletons does not always yield a singleton.

Elements of $T(X \times Y)$ are "joint distributions." One would hope to construct a map Γ: $TX \times TY \to T(X \times Y)$ whose image is the set of "independent" joint distributions realized by "simultaneous" consideration of two distributions. The map $e_{X \times Y}$ of (B) allows this for crisp distributions and systematic use of (C) produces two canditates Γ_1, Γ_2 for Γ, depending on which variable is fixed first. *Commutative theories* are those in which $\Gamma_1 = \Gamma_2$ and these include the fundamental examples.

But a much deeper understanding results from thinking of distributions as operations. Consider $\omega \in T\{1,..., n\}$. In crisp set theory, $\omega \in \{1,..., n\}$ has the operational interpretation that given any n-tuple $(x_1,..., x_n)$ in a set X, ω operates to choose x_ω, that is, ω induces the functions $\hat{\omega}_X$: $X^n \to X$ by $\hat{\omega}_X(x_1,..., x_n) = x_\omega$. In the fuzzy world where $x_1,..., x_n \in TX$ as well as ω are only vague specifications, there is still the induced operation defined by

$$\begin{aligned} (TX)^n &\xrightarrow{\hat{\omega}_X} TX, \\ \alpha &\longmapsto \alpha^{\#}(\omega), \end{aligned} \tag{H}$$

where we write α: $\{1,..., n\} \to TX$ instead of $(x_1,..., x_n)$; thus, $\alpha^{\#}$ as in (C) has form $T\{1,..., n\} \to TX$ which returns an element of TX when evaluated on the fixed ω. Examples are given in Section 3.

We may then say that a function $TX \to TY$ is a *homomorphism* if it commutes with all of the operations, a standard definition in algebra. More generally, say that a function $TX_1 \times \cdots \times TX_n \to TY$ is an *n-homomorphism* if it is a homomorphism in each variable separately. The commutative theories are characterized as those theories admitting the "fuzzification principle" that each f: $X_1 \times \cdots \times X_n \to TY$ has a unique n-homomorphic extension $\bar{f}$: $TX_1 \times \cdots \times TX_n \to TY$. For example, the map Γ: $TX \times TY \to T(X \times Y)$ mentioned above is the unique 2-homomorphic extension of $e_{A \times B}$. Further, any algebraic operation on a set, call it f: $X^n \to X$, induces $e_X f$: $X^n \to TX$ and hence $\overline{e_X f}$: $(TX)^n \to TX$ which is how, in particular, Boolean polynomials $\{\mathbf{true}, \mathbf{false}\}^n \to \{\mathbf{true}, \mathbf{false}\}$ lift to $\mathscr{E}$ as we claimed they did above. Linton [21] and Kock [19] called attention to n-homomorphisms in the context of commutative theories. (Kock was interested primarily in the condition that Γ: $TX \times TY \to T(X \times Y)$ is an isomorphism so that "every joint distribution is independent"; a degenerate condition from our perspective. This condition becomes interesting, however, when the functions in (A, B, C) are allowed to roam over more general closed categories than the category of sets and functions, as they do in Kock's work.)

To make proper contact with work in the logic of computer programs, a preliminary obstacle is to interpret a function of form $X \to T(Y+Z)$ (here $Y+Z$ denotes disjoint union; see flowchart 8.A) as a "conditional statement" **if** p **then** α **else** β for some proposition p: $X \to \mathscr{E}$ and some α: $X \to TY$, β: $X \to TZ$. Two different solutions are presented. The first is based on the Boolean polynomial

$$\begin{array}{llll} \{\mathbf{true}, \mathbf{false}\}^3 & \xrightarrow{\text{if-then-else}} & \{\mathbf{true}, \mathbf{false}\} & \\ (p, f, g) & \longmapsto & f & \text{if } p = \mathbf{true} \\ & & g & \text{if } p = \mathbf{false}. \end{array} \tag{I}$$

(Indeed, all other Boolean operations may be defined in terms of if-then-else—see [25] for a complete discussion—with equations such as

$$\begin{aligned} p \vee q &= \text{if } p \text{ then } \mathbf{true} \text{ else } q, \\ p' &= \text{if } p \text{ then } \mathbf{false} \text{ else } \mathbf{true}, \\ p \wedge q &= (p' \vee q')', \\ p \Rightarrow q &= p' \vee q, \end{aligned} \tag{J}$$

all of which continue to hold in the logic of propositions by the metatheorem already mentioned. On the other hand, the metatheorem does not apply to

$$\text{if } p \text{ then } f \text{ else } g = (p \wedge f) \vee (p' \wedge g). \tag{K}$$

From the fuzzy theories point of view, if-then-else is then seen as more basic than $\vee$, $\wedge$, $(-)'$.) More generally, consider the "Boolean conditional" bc_X: $\{\mathbf{true}, \mathbf{false}\} \times X \times X \to X$ defined by $bc_X(\mathbf{true}, x, y) = x$, $bc_X(\mathbf{false}, x, y) = y$. For commutative theories, the fuzzification principle yields the extension $\widetilde{bc}_X$: $\mathscr{E} \times TX \times TX \to TX$. On the other hand, the operational interpretation of a truth value induces the "distributional conditional" dc_X:

$\mathscr{E} \times TX \times TX \to TX$. Either conditional may be used to define **if** p **then** α **else** β, though dc_X was chosen for the reasons discussed in Section 8. The motivating problem of representing functions in conditional form axiomatizes the "conditional-complete" theories.

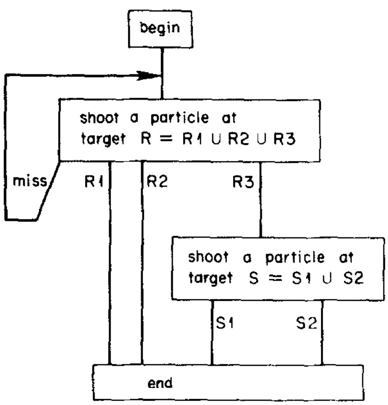
Because fuzzy theories are coextensive with the algebraic theories of universal algebra [24, Exercise 12, p. 32] there is an abundance of examples as well as extensive information on how examples are generated. The issues of importance in this paper are different from those of universal algebra, however. For instance, many of the theorems do not extend nicely to noncommutative theories whereas from a universal algebraic perspective the commutative theories constitute a rarified class of examples whose general theory is not much richer.

Many topics have been omitted. Algebras (including their fuzzy theory interpretation) and generalizations to arbitrary categories are treated in [24]. Automaton realization problems are discussed in [2, 3].

The earliest construct equivalent to the fuzzy theories of this paper are the "standard constructions" in the appendix of [13]. The founding work is [20]. A more complete history is included in the end-of-sections notes of [24].

1. Fuzzy Theories

Consider the algorithm shown in (A).



(A)

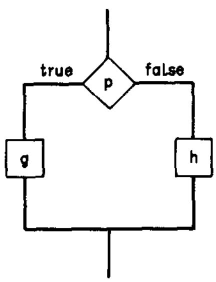
In a "dartboard" scenario, one may imagine that this algorithm is crisp, terminating in exactly one of $\{R1, R2, S1, S2\}$. Another interpretation is provided by the following fragment of the American game of baseball: $R =$ batter with full count, miss $=$ foul ball, $R1 =$ strike 3, $R2 =$ ball 4, $R3 =$ hit ball, $S1 =$ batter out, $S2 =$ batter safe. Traditionally, this algorithm is crisp most of the time but fuzzy a significant number of times. Quantum theory is a source of related algorithms whose outcome is fuzzy "even in principle." The "branched questionnaires" of [29, Section 4] also give rise to fuzzy algorithms of this type.

Modern computer programming languages such as PASCAL emphasize "structured programming" (see [1, and the bibliography therein]) as a systematic tool in the analysis and synthesis of flowcharts and programs. There are three operations from which all flowcharts are to be built from atomic ones:

composition t; u

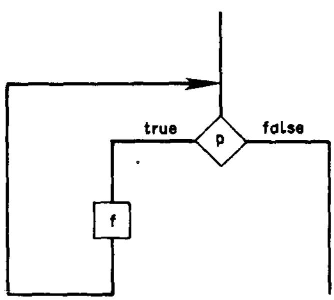
(B)

conditional **if** p **then** g **else** h



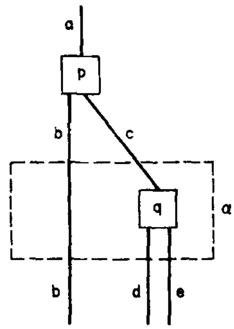
(C)

iteration **while** p **do** f



(D)

A flowchart is *loop-free* if it can be built without using (D). Our definition of fuzzy theory can be motivated entirely by the need to interpret the loop-free scheme



(E)

(It is not hard to adapt the results of [10] to show that fuzzy theories can interpret any loop-free scheme, so we will not belabor that point in this paper.)

DEFINITION 1.1. A *fuzzy theory* is $\mathbf{T} = (T, e, (-)^{\#})$ where,

T assigns to each set X a set TX,

e assigns to each set X a function $e_X : X \to TX$,

$(-)^{\#}$ assigns to each function α: $X \to TY$ a function $\alpha^{\#}$: $TX \to TY$ subject to the following three axioms on arbitrary α: $X \to TY$, β: $Y \to TZ$.

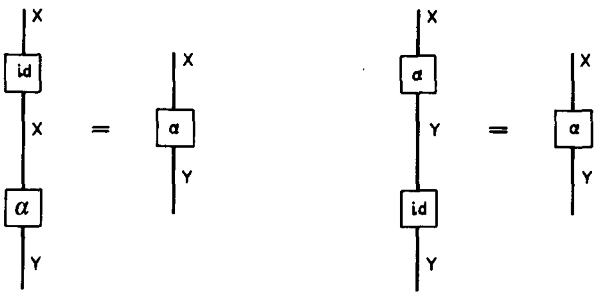
extension axiom $\alpha^{\#} e_X = \alpha$ (where juxtaposition denotes composition),

post-identity axiom $(e_X)^{\#} = id_{TX}$ (where $id_Y : Y \to Y, y \mapsto y$),

associativity axiom $(\beta^{\#}\alpha)^{\#} = \beta^{\#}\alpha^{\#}$.

HEURISTICS 1.2. In a fuzzy interpretation of (E) the "outcome" is a vaguely-specified element of $\{b, d, e\}$, more generally motivating the passage from X to TX as in (0.A). We have not limited the construction to finite X because there are theoretical advantages in being able to form $T(TX)$ whereas TX is often infinite even when X is finite (0.E, F). Now (E) has the form of the composition p; α. Since the output of p is already vague, this motivates the need for $\alpha^{\#}$, the output of (E) being $\alpha^{\#}(p(a))$. Even so, the semantics of (E) is not precise because the data determining (E) take the form p: $\{a\} \to T\{b, c\}$, q: $\{c\} \to T\{d, e\}$ and we have yet to explain how to write down α: $\{b, c\} \to T\{b, d, e\}$. Intuitively, (i) $\alpha(b) = b$ whereas (ii) $\alpha(c) = q(c)$. The obstruction to (i) is overcome by e_X, the more precise description being $\alpha(b) = e_{\{b,d,e\}}(b)$. The heuristic meaning of e_X is "the inclusion of the crisp distributions among the vague ones." While we have not assumed that e_X is injective, this is proved in Theorem 4.4 below. Continuing, the problem with (ii) is that we do not yet know how to think of $T\{d, e\}$ as a subset of $T\{b, d, e\}$. This difficulty is overcome as follows. Whenever X is a subset of Y let $f: X \to Y$ be the inclusion map, $f(x) = x$, and set $\beta: X \to TY = e_Y f$. Then $\beta^{\#}: TX \to TY$ provides the desired map (which is proved injective in Theorem 4.3 below). The more precise version of (ii), then, is $\alpha(c) = (e_{\{b,d,e\}} f)^{\#}(q(c))$, where $f: \{d, e\} \to \{b, d, e\}$ is inclusion.

For more complex schemes than (E) there will be compositions $\alpha;\beta$ in which the first term has more than one input line. The definition is clear: for $\alpha: X \to TY$, $\beta: Y \to TZ$, $\alpha;\beta = \beta^{\#}\alpha: X \to TZ$. The associativity axiom asserts that this composition is associative (proof: for $\gamma: Z \to TW$, if $(\alpha;\beta);\ \gamma = \alpha;\ (\beta;\gamma)$ set $\alpha = id_{TX}$ whence $(\gamma^{\#}\beta)^{\#} = (\gamma^{\#}\beta)^{\#}\alpha = \alpha;\ (\beta;\gamma) = (\alpha;\beta);\ \gamma = \gamma^{\#}(\beta^{\#}\alpha) = \gamma^{\#}\beta^{\#}$ whereas, conversely, $(\alpha;\beta);\ \gamma = \gamma^{\#}(\beta^{\#}\alpha) = (\gamma^{\#}\beta^{\#})\alpha = (\gamma^{\#}\beta)^{\#}\alpha = \alpha;\ (\beta;\gamma)$). The extension and post-identity axioms are motivated, respectively, by the flowchart tautologies



The extension axiom derives its name from the fact that $\alpha^{\#}$ extends α whereas the post-identity axiom is named after its flowchart tautology.

FUNDAMENTAL EXAMPLE 1.3. Crisp Set Theory. $TX = X$, $e_X = id_X$, $\alpha^{\#} = \alpha$.

FUNDAMENTAL EXAMPLE 1.4. Fuzzy Set Theory. Let TX be the set of functions from X to the unit interval $[0, 1]$. Let $e_X(x)$ be the characteristic function of $\{x\}$, that is, $e_X(x)$ maps y to 1 if $y = x$ and maps y to 0 otherwise. If $\alpha: X \to TY$, $x \mapsto \alpha_x: Y \to [0, 1]$, define $\alpha^{\#}: TX \to TY$ by

$$\alpha^{\#}(p): Y \to [0, 1], \qquad y \mapsto \mathrm{Sup}_x \operatorname{Min}(p(x), \alpha_x(y)).$$

The composition $\beta^{\#}\alpha$ is the composition of fuzzy relations of [28, p. 346].

FUNDAMENTAL EXAMPLE 1.5. Probabilistic Set Theory. Define TX to be the set of finite support probability distributions on X. Thus an element of X

is a function p: $X \to [0, 1]$ such that $p(x) = 0$ for all but finitely many x and $\sum p(x) = 1$. Let $e_X(x)$ assign probability 1 to x and hence probability 0 to every other element. If α: $X \to TY$, $x \mapsto \alpha_x$: $Y \to [0, 1]$, define $\alpha^{\#}$: $TX \to TY$ by

$$\alpha^{\#}(p)\colon Y \to [0, 1], \qquad y \mapsto \sum_x p(x)\, \alpha_x(y).$$

Here a function $X \to TY$ amounts to a column-stochastic matrix with X indexing columns and Y indexing rows and composition $\alpha; \beta$ is the usual composition $\beta\alpha$ of column-stochastic matrices.

FUNDAMENTAL EXAMPLE 1.6. Possibilistic Set Theory. TX is the set of subsets of X, $e_X(x) = \{x\}$, $\alpha^{\#}(p) = \bigcup (\alpha(a)\colon a \in p)$. Composition is the usual one for binary relations.

EXAMPLE 1.7. Credibility Theory. In this example, values are unambiguous but their accuracy is vague. Let C be a partially ordered set of "credibility values" possessing binary infima and a greatest element 1. Let $TX = C \times X$. Define $e_X(x) = (1, x)$. Given α: $X \to TY$, α decomposes $\alpha(x) = (\rho(x), f(x))$ into a proviso function ρ and a value function f, that is, "$\alpha(x) = f(x)$ with credibility $\rho(x)$." Define $\alpha^{\#}$ by $\alpha^{\#}(c, x) = (\mathrm{Min}(c, \rho(x)), f(x))$. Then composition is described by "if $\alpha(x) = y$ with credibility c_1 and $\beta(y) = z$ with credibility c_2 then $(\alpha; \beta)(x) = z$ with credibility $\mathrm{Min}(c_1, c_2)$."

EXAMPLE 1.8. Priority Theory. Let TX be the set of all strings $x_1 \cdots x_n$ with $n \geqslant 1$, each $x_i \in X$ and with no repetitions, that is, $x_i \neq x_j$ if $i \neq j$. A distribution is a "choice of outcomes in order of priority" with unmentioned elements of X "abstentions." Define $e_X(x) = x$. For α: $X \to TY$ define $\alpha^{\#}(x_1 \cdots x_n)$ to be the string obtained from the juxtaposition $\alpha(x_1) \cdots \alpha(x_n)$ by keeping the leftmost occurrence of each symbol and deleting the others. (Thus if $\alpha(x_1) = y_1 y_2$ and $\alpha(x_2) = y_1 y_3 y_2$, $\alpha^{\#}(x_1 x_2) = y_2 y_2 y_3$.)

EXAMPLE 1.9. Neighborhood Theory. If X is a topological space, the neighborhood filter $\mathcal{N}$ of $x \in X$ is a family of subsets of X possessing the following four properties:

(i) $X \in \mathcal{N}$,

(ii) If $N_1, N_2 \in \mathcal{N}$ then $N_1 \cap N_2 \in \mathcal{N}$,

(iii) If $N \subset S$ and $N \in \mathcal{N}$ then $S \in \mathcal{N}$,

(iv) $\bigcap \mathcal{N} \neq \varnothing$.

For any set X, let TX be the set of all families $\mathcal{N}$ of subsets of X possessing these four properties. Define $e_X(x) = \{A \subset X\colon x \in A\}$. For α: $X \to TY$, define $\alpha^{\#}(\mathcal{N}) = \{B \subset Y\colon \{x \in X\colon B \in \alpha(x)\} \in \mathcal{N}\}$.

We conclude this section with a brief treatment of fuzzy theories as algebraic objects, defining homomorphisms between theories (which we call theory maps), quotient theories, subtheories and product theories.

DEFINITION 1.10. Let $\mathbf{T} = (T, e, (-)^{\#})$, $\bar{\mathbf{T}} = (\bar{T}, \bar{e}, (-)^{\#\#})$ be two fuzzy theories. A *theory map* λ: $\mathbf{T} \to \bar{\mathbf{T}}$ assigns a function λ_X: $TX \to \bar{T}X$ to each set X, subject to two axioms. The first is (F) which asserts that $\lambda_X e_X(x) = \bar{e}_X$ for

$$\begin{tikzcd} & TX \arrow[d, "\lambda_X"] \\ X \arrow[ur, "e_X"] \arrow[r, "\bar{e}_X"'] & \bar{T}X \end{tikzcd} \tag{F}$$

all X, $x \in X$. The second axiom asserts that given α: $X \to TY$ and $\bar{\alpha}$ defined

$$\begin{tikzcd} & TY \arrow[d, "\lambda_Y"] \\ X \arrow[ur, "\alpha"] \arrow[r, "\bar{\alpha}"'] & \bar{T}Y \end{tikzcd} \tag{G}$$

by (G), then the commutative square (H) obtains. It is trivial to verify that

$$\begin{tikzcd} TX \arrow[r, "\alpha^{\#}"] \arrow[d, "\lambda_X"'] & TY \arrow[d, "\lambda_Y"] \\ \bar{T}X \arrow[r, "\bar{\alpha}^{\#\#}"'] & \bar{T}Y \end{tikzcd} \tag{H}$$

id_X: $TX \to TX$ always defines a theory map $\mathbf{T} \to \mathbf{T}$ and that if λ: $\mathbf{T} \to \mathbf{S}$ and μ: $\mathbf{S} \to \mathbf{R}$ are theory maps then so is $\mu\lambda$: $\mathbf{T} \to \mathbf{R}$ defined by $(\mu\lambda)_X = \mu_X \lambda_X$.

EXAMPLE 1.11. Let $\mathbf{T}$ be priority theory, $\bar{\mathbf{T}}$ be possibility theory. Then λ: $\mathbf{T} \to \bar{\mathbf{T}}$ defined by $\lambda_X(x_1 \cdots x_n) = \{x_1, \ldots, x_n\}$ is a theory map.

DEFINITION 1.12. A *quotient theory of* $\mathbf{T}$ is a theory map λ: $\mathbf{T} \to \bar{\mathbf{T}}$ such that each λ_X is surjective (= onto $\bar{T}X$). Given a theory $\mathbf{T}$ and surjective functions of form λ_X: $TX \to \bar{T}X$ there is at most one way to define $\bar{e}$ and $(-)^{\#\#}$ so as to make $\bar{T}$ into a theory in such a way that λ is a theory map (proof: (F) defines $\bar{e}$ outright; similarly, given $\bar{\alpha}$: $X \to \bar{T}Y$, since λ_Y is surjective there exists a choice function $\alpha\colon X \to TY$ such that (G) holds so that (H) must hold for the chosen α and this is possible for at most one $\bar{\alpha}^{\#\#}$ since λ_X is surjective). Thus if $\mathbf{T}$ is a theory and E_X is an equivalence relation on TX for each set X, the resulting quotient sets $\lambda_X\colon TX \to TX/E_X$ collectively "is or is not" a quotient theory of $\mathbf{T}$.

EXAMPLE 1.13. In fuzzy set theory, $[0, 1]$ is often projected onto $\{0, 1\}$ by choosing a cut point. This idea is tantamount to a theory map as follows. Fix a cut point $0 \leqslant c < 1$, let $\mathbf{T}$ be fuzzy set theory and let $\bar{\mathbf{T}}$ be possibility theory. Then

$$TX \xrightarrow{\lambda_X} \bar{T}X,$$

$$X \xrightarrow{p} [0, 1] \mapsto \{x \in X\colon p(x) > c\}$$

is a quotient theory on $\mathbf{T}$. The verification rests on the following property of the unit interval: if $\operatorname{Sup}(x_i) > c$ then some $x_i > c$. This axiom has been emphasized by Scott [26, p. 110] in a different context.

DEFINITION 1.14. A *subtheory of* $\bar{\mathbf{T}}$ is a theory map $\lambda\colon \mathbf{T} \to \bar{\mathbf{T}}$ such that each λ_X is injective (= one-to-one into $\bar{T}X$). Given a theory $\bar{\mathbf{T}}$ and injective functions of form $\lambda_X\colon TX \to \bar{T}X$ there is at most one way to define e and $(-)^{\#}$ so as to make T into a theory in such a way that λ is a theory map (proof: (F) is possible if and only if $\bar{e}_X$ maps into the image of λ_X in which case $e_X = \lambda_X^{-1}\bar{e}_X$ is the only possible definition; given α, $\bar{\alpha}$ as in (G), $\alpha^{\#}$ in (H) exists if and only if $\bar{\alpha}^{\#\#}\lambda_X$ maps into the image of λ_Y and then $\alpha^{\#} = \lambda_Y^{-1}\bar{\alpha}^{\#\#}\lambda_X$ is the only possible definition). In particular, if TX is defined as a subset of $\bar{T}X$, there is at most one way to make T into a theory such that the inclusions $TX \to \bar{T}X$ constitute a theory map and, if so, we say $\mathbf{T}$ is a *canonical subtheory* of $\bar{\mathbf{T}}$. It is trivial to verify that any intersection of canonical subtheories of $\bar{\mathbf{T}}$ is again a canonical subtheory of $\bar{\mathbf{T}}$ so that, in fact, any construction defining a subset TX of $\bar{T}X$ must generate a canonical subtheory.

EXAMPLE 1.15. $\{p \in [0, 1]^X\colon p(x) = 1$ for some $x\}$ is a canonical subtheory of fuzzy set theory (cf. the "normalized fuzzy variables" of [11, p. 180]).

EXAMPLE 1.16. Let $\mathbf{T}$ be credibility theory with credibility poset $[0, 1]$ and let $\bar{T}$ be fuzzy set theory. Then $\lambda_X(c, x)(y) = c$ if $x = y$, $=0$ if $x \neq y$ is a subtheory $\mathbf{T} \to \bar{\mathbf{T}}$.

EXAMPLE 1.17. "Non-empty" and "finite" define subtheories of possibility theory. Non-empty possibility theory is a subtheory of neighborhood theory via $\lambda_X(A) = \{B \subset X: \ A \subset B\}$. Finite non-empty possibility theory is *not* a subtheory of probabilistic set theory if $\lambda_X\{x_1,..., x_n\}$ assigns probability $1/n$ to each x_i. In this precise sense, a set of possibilities is not a set of equally likely outcomes.

DEFINITION 1.18. If $(\mathbf{T}_i: \ i \in I)$ is a family of fuzzy theories, their *product* $\prod \mathbf{T}_i$ is the theory $\mathbf{T}$ defined by

$$TX = \prod T_i X_i,$$

$$e_X(x_i : i \in I) = ((e_i)_X(x_i): i \in I).$$

Given $\alpha: X \to TY$ write $\alpha(x) = (\alpha_i(x): i \in I)$; then for $p = (p_i : i \in I) \in TX$, $\alpha^\#(p) = (\alpha_i^\#(p_i): i \in I)$.

DEFINITION 1.19. A theory map λ: $\mathbf{T} \to \bar{\mathbf{T}}$ is an *isomorphism* if each λ_X is surjective and injective. It is not hard to see that, in this case, λ_X^{-1} constitutes a theory map. $\mathbf{T}$ and $\bar{\mathbf{T}}$ are *isomorphic* if there exists an isomorphism from $\mathbf{T}$ and $\bar{\mathbf{T}}$. For example, $\{p \in [0, 1]^X: \ p(x) \in \{0, 1\}$ for all $x\}$ describes a subtheory of fuzzy set theory which is isomorphic to possibilisic set theory.

2. EQUALITY AND DEGREE OF MEMBERSHIP

For the remainder of the paper mention of $\mathbf{T} = (T, e, (-)^\#)$ without further modification refers to an arbitrary fuzzy theory.

In this section we define fuzzy truth values, define the equality of distributions as a fuzzy truth valued function and explore properties of this equality.

DEFINITION 2.1. Truth Values. Fix a two-element set $\{$**true, false**$\}$ of crisp truth values and denote this set simply as 2. Define the set of $\mathbf{T}$-*truth values* as the set $\mathscr{T} = T2$ of $\mathbf{T}$-distributions on **true** and **false.** We will generally write **true** $\in \mathscr{T}$ for the more cumbersome e_2(**true**) and similarly for **false**.

DEFINITION 2.2. The Degree-of-Membership Map. We regard 2^X as both the set of subsets of X and the set of functions from X to 2 via the usual identification $S \mapsto \chi_S$, where $\chi_S(x) = \mathbf{true}$ if $x \in S$ and $= \mathbf{false}$ if $x \notin S$. For each $S \subset X$, define $T\chi_S : TX \to \mathscr{T}$ (see (A)) by $T\chi_S = (e_2\chi_S)^{\#}$. We then define

$$\begin{tikzcd} X \arrow[r, "\chi_S"] \arrow[d, "e_X"'] & 2 \arrow[d, "e_2"] \\ TX \arrow[r, "T\chi_S"'] & \mathscr{T} \end{tikzcd} \tag{A}$$

the *degree-of-membership maps*

$$2^X \times TX \xrightarrow{dm_X} \mathscr{T}$$

by $dm_X(S, p) = (T\chi_S)(p)$. We call $T\chi_S = dm_X(S, -)$ the **T**-*characteristic function* of S. We also use the same notation dm_X: $X \times TX \to \mathscr{T}$ for the restriction to singletons, $dm_X(x, p) = dm_X(\{x\}, p)$.

OBSERVATION 2.3. The Exponential Law of Set Theory. There is a bijective correspondence

$$\frac{X \times Y \xrightarrow{f} Z}{Y \xrightarrow[g]{} Z^X} \tag{B}$$

given by $g(y)(x) = f(x, y)$. Similarly,

$$\frac{X \times Y \xrightarrow{f} Z}{X \xrightarrow[h]{} Z^Y}. \tag{C}$$

OBSERVATION 2.4. Coordinatewise Extension. Given α: $X \to (TX)^Z$, define $\alpha^{\#}$: $TX \to (TY)^Z$ by $\alpha^{\#}(x) = ((\alpha_z)^{\#}(x): z \in Z)$, where $\alpha_z : X \to TY$ is α composed with the z-coordinate projection $(TY)^Z \to TY$.

DEFINITION 2.5. The Equality Map. The **T**-*equality maps*

$$TX \times TX \xrightarrow{eq_X} \mathscr{T}$$

(with the interpretation "$p = q$ has truth value $eq_X(p, q)$") are defined as

Step 1. Begin with the degree-of-membership map restricted to singletons

$$X \times TX \xrightarrow{dm_X} \mathscr{T}.$$

Step 2. Apply exponential law (C) to Step 1:

$$X \longrightarrow \mathscr{T}^{TX}.$$

Step 3. Apply coordinatewise extension to Step 2:

$$TX \longrightarrow \mathscr{T}^{TX}.$$

Step 4. Apply exponential law (C) to Step 3:

$$TX \times TX \xrightarrow{eq_X} \mathscr{T}$$

Note: By construction, $eq_X(e_X(x), p) = dm_X(x, p)$.

EXAMPLE 2.6. Crisp Equality. Let **T** be crisp set theory. Then $\mathscr{T} = 2$, $T\chi_S = \chi_S$ and $eq_X(p, q) = \mathbf{true}$ if $p = q$ and $= \mathbf{false}$ if $p \neq q$.

EXAMPLE 2.7. Equality of Fuzzy Sets. Let **T** be fuzzy set theory. The set of truth values is *not* the unit interval $[0, 1]$ but is, rather, $[0, 1]^2$ whose typical element $t = (t_{\mathbf{true}}, t_{\mathbf{false}})$ consists of "a degree of truth and a degree of falsity." This idea was also suggested by [11, p. 180]. Given $p, q\colon X \to [0, 1]$,

$$\begin{aligned} eq_X(p, q) = t, \quad t_{\mathbf{true}} &= \sup_x \operatorname{Min}(p(x), q(x)) \\ t_{\mathbf{false}} &= \operatorname*{Sup}_{x \neq y} \operatorname{Min}(p(x), q(y)). \end{aligned} \tag{D}$$

In particular, the usual "degree-of-membership of x in p" in the fuzzy set literature, namely, $p(x)$, is the **true** coordinate of $dm_X(x, p)$ which, however, also has **false** coordinate $\operatorname{Sup}_{y \neq x} p(y)$. In general, $dm_X(S, p)$ has **true** coordinate $\operatorname{Sup}(p(x)\colon x \in S)$ and has **false** coordinate $\operatorname{Sup}(p(x)\colon x \notin S)$.

EXAMPLE 2.8. Probabilistic Equality. Let **T** be probabilistic set theory. Then $\mathscr{T}$ may be identified with the unit interval under the bijection $t \leftrightarrow$ probability of **true**. Then

$$eq_X(p, q) = \sum p(x)\, q(x) \tag{E}$$

a familiar formula for the probability of equality of two independent random variables on a finite probability space. Notice that $dm_X(x, p)$ is just $p(x)$ and that, in general, $dm_X(S, p) = \sum (p(x)\colon x \in S)$.

EXAMPLE 2.9. Possibilistic Equality. Let **T** be possibilistic set theory. Then $\mathscr{T}$ has four elements $\varnothing$, {**false**}, {**true**}, {**false, true**} which we shall respectively relabel as **undefined, no, yes** and **maybe**. Then

$$eq_X(p, q) = \textbf{undefined} \quad \text{if } p \text{ is empty or } q \text{ is empty}$$
$$= \textbf{no} \quad \text{if } p, q \text{ are nonempty and disjoint} \tag{F}$$
$$= \textbf{yes} \quad \text{if } p = \{x\} = q \text{ for some } x$$
$$= \textbf{maybe} \quad \text{else.}$$

Also, $dm_X(S, p) = \textbf{yes}$ if p is a nonempty subset of S and is $eq_X(S, p)$ in every other case.

EXAMPLE 2.10. Equality for the Credibility Theory. For **T** as in 1.7, $eq_X((c, x), (c', x')) = \textbf{true}$ with credibility $\mathrm{Min}(c, c')$ if $x = x'$, and $=\textbf{false}$ with credibility $\mathrm{Min}(c, c')$ if $x \neq x'$.

EXAMPLE 2.11. Equality for the Priority Theory. Let **T** be as in 1.8. Then $\mathscr{E}$ has four elements **true, false, true false, false true** which we shall respectively relabel **true, false, moretruethanfalse, morefalsethantrue**. For $p = p_1 \cdots p_m$, $q = q_1 \cdots q_n$, equality is given by

$$eq_X(p, q) = \textbf{true} \quad \text{if } p = q = p_1$$
$$= \textbf{false} \quad \text{if } p_1 \neq q_j \text{ for all } i, j \tag{G}$$
$$= \textbf{moretruethanfalse} \quad \text{if } p_1 = q_1, \text{ some } p_i \neq q_j$$
$$= \textbf{morefalsethantrue} \quad \text{if } p_1 \neq q_1, \text{ some } p_i = q_j.$$

EXAMPLE 2.12. Equality for the Neighborhood Theory. In general, equality relative to a subtheory is computed in the ambient theory. In particular, equality relative to the "non-empty" subtheory of possibilistic set theory is just as in (F) save that **undefined** is deleted from $\mathscr{E}$ and the first case in (F) should be deleted. Let **T** be the theory of 1.9. If λ represents non-empty possibility theory as a subtheory of **T** as in 1.17, λ_X is bijective when X is finite (because if $\mathscr{N} \in TX$, $\mathscr{N}$ is finite so that $\bigcap \mathscr{N} \in \mathscr{N}$). Thus for **T**, $\mathscr{E} = \{\textbf{yes, no, maybe}\}$ and (F) describes eq_X when X is finite. For general X, some terminology is helpful. Given $\mathscr{N} \in TX$, $x \in X$ say that $\mathscr{N}$ *converges to* x if $\{x\} \in \mathscr{N}$ and say that $\mathscr{N}$ *excludes* x if there exists $N \in \mathscr{N}$ with $x \notin N$. Then

$$eq_X(\mathscr{N}, \mathscr{M}) = \textbf{yes} \quad \text{if } \mathscr{N}, \mathscr{M} \text{ converge to a common point}$$
$$= \textbf{no} \quad \text{if } \{x : \mathscr{M} \text{ excludes } x\} \in \mathscr{N} \tag{H}$$
$$= \textbf{maybe} \quad \text{else.}$$

To clarify the first case, notice that if $\mathscr{N}$ converges to x, $\mathscr{N} = \{A : x \in A\}$. In general, $dm_X(S, \mathscr{N})$ is **yes** if S belongs to $\mathscr{N}$, is **no** if the complement of S belongs to $\mathscr{N}$ and is otherwise **maybe**.

DEFINITION 2.13. Anti-Reflexive Theories. As is reenforced by the following table, $eq_X(p, p)$ is the "degree of vagueness" of $p \in TX$.

Theory	$eq_X(p, p)$
Crisp set theory	**true**
Fuzzy set theory	$\mathrm{Sup}_x\, p(x)$, $\mathrm{Sup}_x(\mathrm{Min}(p(x), \mathrm{Sup}_{y \neq x}\, p(y)))$ (**true** coordinate first)
Probabilistic set theory	$\sum p(x)\, p(x)$
Possibilistic set theory	**undefined, yes, maybe** accordingly as p is empty, crisp, otherwise
Credibility theory	$(c, \mathbf{true})$ if $p = (c, x)$
Priority theory	**true** if p is crisp, **moretruethanfalse** else
Neighborhood theory	**true** if p is crisp, **maybe** else

$\mathbf{T}$ is *anti-reflexive* if for X, $p \in TX$, if $eq_X(p, p) = \mathbf{true}$ then p is crisp. All of the theories in the table above are anti-reflexive. Any subtheory of an anti-reflexive theory is anti-reflexive and any product of anti-reflexive theories is anti-reflexive.

EXAMPLE 2.14. A Theory Which is Not Anti-Reflexive. The credibility poset of 1.7 is a monoid with infimum as multiplication and greatest element as unit. More generally, if C is any monoid, the construction of 1.7 with multiplication replacing infimum and with unit replacing greatest element produces a fuzzy theory. The formula of 2.10 generalizes, and if $p = (c, x)$, $eq_X(p, p) = (cc, \mathbf{true})$. The equation "$eq_X(p, p) = \mathbf{true}$" here, then, is "$(cc, \mathbf{true}) = (1, \mathbf{true})$" which amounts to the requirement that $cc = 1$. Now, for example, take M the monoid of subsets of a set with symmetric difference as multiplication and empty set as unit. In this example $cc = 1$ holds for every c, so that $eq_X(p, p) = \mathbf{true}$ for every p.

DEFINITION 2.15. Symmetry of Equality. Say that $\mathbf{T}$*-equality is symmetric* if for every X and for every $p, q \in TX$, $eq_X(p, q) = eq_X(q, p)$. This condition holds for every theory mentioned so far except the neighborhood theory. Indeed, let $\mathbf{T}$ be the neighborhood theory, let X be the real line, and let $\mathscr{N}_x$ be the filter of neighborhoods of $x \in X$ in the usual topology. On the one hand, it is true that $eq_X(\mathscr{N}_x, \mathscr{N}_y)$ is **maybe** when $x = y$ and **no** when $x \neq y$. On the other hand, fix x and set $M = \{y \in X\colon\ y \neq x\}$, $\mathscr{M} = \{M, X\} \in TX$. Then $eq_X(\mathscr{M}, \mathscr{N}_x) = \mathbf{maybe}$ whereas $eq_X(\mathscr{N}_x, \mathscr{M}) = \mathbf{no}$.

DEFINITION 2.16. The Eigenstate Condition. In quantum mechanics, the act of a measuring an observable forces a crisp state. While the analogy is loose, it suggests the following colorful terminology. A theory **T** satisfies the *eigenstate condition* if for all X, and for all $x \in X, p \in TX$, if $dm_X(x, p) = \mathbf{true}$ then $p = e_X(x)$. All of the examples considered so far satisfy this condition.

EXAMPLE 2.17. A Theory Not Satisfying the Eigenstate Condition. Define a modification of priority theory as follows. TX is the set of all repetition-free strings $x_1 \cdots x_n$ $(n \geqslant 0)$ in which the empty string Λ is now allowed. Define $e_X(x) = x$, and obtain $\alpha^{\#}(x_1 \cdots x_n)$ from $\alpha(x_1) \cdots \alpha(x_n)$ by deleting repeated symbols. For example $\alpha^{\#}(\Lambda) = \Lambda$, and for $\alpha(x_1) = y_1 y_2$, $\alpha(x_2) = y_3 y_2 y_4$, $\alpha^{\#}(x_1 x_2) = y_1 y_3 y_4$. The set $\mathscr{E}$ of truth values consists of those of 2.11 together with Λ. To see that the eigenstate condition fails, observe that $dm_X(x, xyz) = \mathbf{true}$ (because **true false false** reduces to **true**).

DEFINITION 2.18. Faithful Theories. The set of **T**-*propositions on* X is defined to be the set $\mathscr{E}^X$ of all functions from X to $\mathscr{E}$. **T** is *faithful* if for all X the *representation map*

$$\begin{aligned} TX &\longrightarrow \mathscr{E}^X, \\ p &\longmapsto dm_X(-, p) \end{aligned} \tag{I}$$

is injective. The four fundamental examples are faithful.

DEFINITION 2.19. Propositional Completeness. **T** is *propositionally complete* if for all X, whenever p, q are distinct elements of TX there exists a proposition α: $X \to \mathscr{E}$ with $\alpha^{\#}(p) \neq \alpha^{\#}(q)$. The four fundamental examples are propositionally complete because of

THEOREM 2.20. *A faithful theory is propositionally complete.*

Proof. Just observe that $dm_X(x, p) = \alpha^{\#}(p)$ for $\alpha = e_2 \chi_{\{x\}}$. ∎

EXAMPLE 2.21. The converse of 2.20 Fails. The neighborhood theory is propositionally complete. To see this, if $\mathscr{N} \neq \mathscr{M}$ then there exists, say, $N \in \mathscr{N}$ with $N \notin \mathscr{M}$. Define $\alpha = \chi_N$: $X \to \mathscr{E}$. Then $\alpha^{\#}(\mathscr{N}) = dm_X(N, \mathscr{N}) = \mathbf{yes}$ whereas $\alpha^{\#}(\mathscr{M}) = dm_X(N, \mathscr{M}) \neq \mathbf{yes}$. On the other hand, this theory is not faithful since it is not hard to show that the cardinal of TX for infinite X is larger than the cardinal of $\mathscr{E}^X$.

EXAMPLE 2.22. The Priority Theory is Not Propositionally Complete. Indeed, if x, y, z are distinct elements of X, xyz and $xzy \in TX$ cannot be distinguished by any proposition.

3. Distributions as Operations

In this section we show that distributions may be equivalently viewed as operations.

Definition 3.1. The Operation Induced by a Distribution. Let n be a set. (Despite the notation, n is not the special case $\{0,..., n-1\}$; n is any set). Fix $\omega \in Tn$. For each X, ω induces a function of form $\hat{\omega}_X: (TX)^n \to TX$ defined by

$$\hat{\omega}_X(\alpha) = \alpha^\#(\omega) \tag{A}$$

for each n-tuple of distributions $\alpha: n \to TX \in (TX)^n$.

Definition 3.2. Abstract Operations. Let n be a set. An *abstract n-ary* **T**-*operation* τ assigns to each set X a function of form $\tau_X: (TX)^n \to TX$ subject to the coherence requirement that (B) holds for every $\beta: X \to TY$.

$$\begin{tikzcd} (TX)^n \arrow[r, "\tau_X"] \arrow[d, "(\beta^\#)^n"'] & TX \arrow[d, "\beta^\#"] \\ (TY)^n \arrow[r, "\tau_Y"'] & TY \end{tikzcd} \tag{B}$$

Diagram (B) asserts that for each n-tuple $\alpha: n \to TX$, $\tau_Y(\beta^\#\alpha) = \beta^\#\tau_X(\alpha)$. In terms of the composition $\alpha; \beta = \beta^\#\alpha$, the condition is simply that τ respects composition: $\tau_Y(\alpha; \beta) = (\tau_X(\alpha)); \beta$.

Theorem 3.3. *The passage* $\omega \mapsto \hat{\omega}$ *of* 3.1 *establishes a bijection from the set Tn of distributions on n to the set of abstract n-ary* **T**-*operations.*

Proof. To prove that (B) commutes for $\hat{\omega}$ use the associativity axiom for **T**: $\hat{\omega}_Y(\beta^\#\alpha) = (\beta^\#\alpha)^\#(\omega) = \beta^\#\alpha^\#(\omega) = \beta^\#\hat{\omega}_X(\alpha)$. Now let τ be an arbitrary abstract n-ary **T**-operation. Each $\alpha: n \to TX$ induces (C) as a special of (B).

$$\begin{tikzcd} (Tn)^n \arrow[r, "\tau_n"] \arrow[d, "(\alpha^\#)^n"'] & Tn \arrow[d, "\alpha^\#"] \\ (TX)^n \arrow[r, "\tau_X"'] & TX \end{tikzcd} \tag{C}$$

Hence if ω is defined to be $\tau_n(e_n) \in Tn$, $\tau = \hat{\omega}$ because using the extension axiom and (C) we have $\tau_X(\alpha) = \tau_X(\alpha^\# e_n) = \alpha^\#(\tau_n(e_n)) = \hat{\omega}_X(\alpha)$. So far, then, we have seen that $\omega \mapsto \hat{\omega}$ is well defined and surjective. To complete the proof we must show that $\hat{\omega}$ is determined by ω and this follows from the post-identity axiom since $\omega = e_n^{\#}(\omega) = \hat{\omega}_n(e_n)$. ∎

EXAMPLE 3.4. Crisp Operations. For $i \in n$, $(e_n(i))_X\colon (TX)^n \to TX$ is the i-coordinate projection as is immediate from the extension axiom. These are the only operations in crisp set theory.

EXAMPLE 3.5. Fuzzy Set Operations. Given $\omega\colon n \to [0, 1] \in Tn$

$$\begin{aligned} ([0,1]^X)^n &\xrightarrow{\hat{\omega}_X} [0,1]^X, \\ (f_i : i \in n) &\longmapsto X \longrightarrow [0,1] \\ & \qquad\qquad x \longmapsto \operatorname*{Sup}_i \operatorname{Min}(\omega(i), f_i(x)). \end{aligned}$$

EXAMPLE 3.6. Probabilistic Operations. The operations in probabilistic set theory are those of convex combination. If $\omega \in Tn$ and if $p_i \in TX$ $(i \in n)$ then $\hat{\omega}_X(p_i) = \sum \omega(i)\, p_i$.

EXAMPLE 3.7. Possibilistic Operations. If $\omega \subset n$ and $p_i \subset X$ $(i \in n)$ then $\hat{\omega}_X(p_i) = \bigcup\, (p_i : i \in n)$.

EXAMPLE 3.8. Operations for the Credibility Theory. If $\omega = (c, j) \in Tn$ and $(c_i, x_i) \in TX$ $(i \in n)$ then $\hat{\omega}_X(c_i, x_i) = (\operatorname{Min}(c, c_j), x_j)$.

EXAMPLE 3.9. Operations for the Priority Theory. If $\omega = i_1 \cdots i_k \in Tn$ and $p_i \in TX$ $(i \in n)$ then $\hat{\omega}_X(p_i)$ is obtained from $p_{i_1} \cdots p_{i_k}$ by keeping the leftmost occurrence of each symbol and deleting all other occurrences.

EXAMPLE 3.10. Operations for the Neighborhood Theory. If $\omega \in Tn$ and if $\mathcal{N}_i \in TX$ $(i \in n)$, $\hat{\omega}_X(\mathcal{N}_i) = \{A \subset X\colon \{i \in n\colon A \in \mathcal{N}_i\} \in \omega\}$.

4. HOMOMORPHISMS

Homomorphisms are maps which respect the T-operations. Each function $f\colon X \to Y$ induces a homomorphism $Tf\colon TX \to TY$ which is injective when f is and which is bijective when f is. Homomorphisms play a role in developing further basic properties of theories in this section.

DEFINITION AND THEOREM 4.1. Let $\varphi\colon TX \to TY$ be a function. The following three conditions on φ are equivalent and define when φ is a *homomorphism*.

1. φ commutes with all **T**-operations, that is, for every abstract n-ary

$$\begin{array}{ccc} (TX)^n & \xrightarrow{\tau_X} & TX \\ {\scriptstyle\varphi^n}\downarrow & & \downarrow{\scriptstyle\varphi} \\ (TY)^n & \xrightarrow[\tau_Y]{} & TY \end{array} \tag{A}$$

T-operation τ, square (A) commutes: $\varphi\tau_X(p_i) = \tau_Y(\varphi p_i)$ for all n-tuples $(p_i : i \in n)$ in TX.

2. $\varphi = (\varphi e_X)^{\#}$.
3. $\varphi = \alpha^{\#}$ for some $\alpha: X \to TY$.

Proof. Condition 1 implies 2. Let $q \in TX$ and consider (A) with $n = X$, $\tau = \hat{q}$. Using (3.A) and the post-identity axiom, $\varphi(q) = \varphi(e_X^{\#}(q)) = \varphi(\hat{q}_X(e_X)) = \hat{q}_Y(\varphi e_X) = (\varphi e_X)^{\#}(q)$ so that $\varphi = (\varphi e_X)^{\#}$.

Condition 2 implies 3. Set $\alpha = \varphi e_X$.

Condition 3 implies 4. This is immediate from the definition (3.B). ∎

OBSERVATION 4.2. Functoriality of T. Given f: $X \to Y$ there is an induced homomorphism Tf: $TX \to TY$ defined by $Tf = (e_Y f)^{\#}$. Then $T(id_X) = e_X^{\#} = id_{TX}$ and for g: $Y \to Z$, $T(gf) = (e_Z gf)^{\#} = (((e_Z g)^{\#} e_Y) f)^{\#} = ((e_Z g)^{\#}(e_Y f))^{\#} = (e_Z g)^{\#}(e_Y f)^{\#} = (Tg)(Tf)$. These two equations—whose verification required all three fuzzy theory axioms—assert that T is a functor from the category of sets to itself. The commutative

$$\begin{array}{ccc} X & \xrightarrow{e_X} & TX \\ {\scriptstyle f}\downarrow & & \downarrow{\scriptstyle Tf} \\ Y & \xrightarrow[e_Y]{} & TY \end{array} \tag{B}$$

square (B) then asserts that e is a natural transformation from the identity functor of the category of sets to T. (For facts about functors and natural transformations see [4, 22].

Note: The **T**-characteristic function of S of 2.2 is an example of this construction.

The second statement in the next result guarantees that "TX is abstract."

THEOREM 4.3. *Let f: $X \to Y$. Then if f is injective, Tf is injective. If f is bijective, Tf is bijective.*

Proof. The second statement depends only on the functoriality of T. For if f is bijective it has an inverse g and $(Tg)(Tf) = T(gf) = T(id_X) = id_{TX}$, $(Tf)(Tg) = id_{TY}$ similarly, so that Tg is inverse to Tf. A similar argument almost proves the first statement. If f is injective and X is not empty then there exists g: $Y \to X$ with $gf = id_X$; since $(Tg)(Tf) = id_{TX}$, Tf is injective. A different argument must be used if X is empty. (We regard the unique function $\varnothing \to X$, "the inclusion of the empty subset," as being injective since if a map is not injective there are two distinct elements in its domain which are mapped to the same element.) There is no problem if $T\varnothing = \varnothing$. Otherwise, there exists a function α: $Y \to T\varnothing$ and hence the homomorphism $\alpha^{\#}$: $TY \to T\varnothing$. Now it is obvious from 4.1 and the associativity axiom that

$$id_{TA} : TA \to TA \text{ is a homomorphism.} \tag{C}$$

$$\begin{aligned} &\text{If } \varphi: TA \to TB, \quad \psi: TB \to TC \text{ are homomorphisms,} \\ &\psi\varphi: TA \to TC \text{ is again a homomorphism.} \end{aligned} \tag{D}$$

Hence $\alpha^{\#}(Tf)$ and $id_{T\varnothing}$: $T\varnothing \to T\varnothing$ are both homomorphisms whereas it is clear from 4.1.2 that there is only one homomorphism $T\varnothing \to T\varnothing$. It follows that $\alpha^{\#}(Tf) = id_{T\varnothing}$ and Tf is injective. ∎

DEFINITION AND THEOREM 4.4. The following four conditions on a fuzzy theory are equivalent and define the class of *consistent* theories.

1. "**true** $\neq$ **false**," that is, $e_2 : 2 \to T2$ is injective.
2. If $f \neq g$: $X \to Y$, then $Tf \neq Tg$: $TX \to TY$.
3. There exists Y such that TY has at least two elements.
4. For all sets X, $e_X : X \to TX$ is injective.

Proof. Condition 1 implies 2. If $f \neq g$: $X \to Y$ there exists h: $Y \to 2$ with $hf \neq hg$ (e.g., if $f(x_0) \neq g(x_0)$ let $h(f(x_0)) =$ **true**, $h(y) =$ **false** for all $y \neq f(x_0)$). Applying (B) twice, we have $T(hf)e_X = e_2 hf$ and $T(hg)e_X = e_2 hg$. As e_2 is injective and $hf \neq hg$ we must have $T(hf)e_X \neq T(hg)e_X$. Applying functoriality, $(Th)(Tf)e_X \neq (Th)(Tg)e_X$ and $Tf \neq Tg$ in particular.

Condition 2 implies 3. This is obvious since there is at most one function $TX \to TY$ if TY has at most one element.

Condition 3 implies 4. If TY has at least two elements, cartesian powers of TY get arbitrarily large and given any set X there exists a set Z and an injection α: $X \to (TY)^Z$. Let $\alpha^{\#}$: $TX \to (TY)^Z$ be the coordinatewise extension of 2.4. Then $\alpha^{\#}e_X = \alpha$. But then e_X is injective because α is.

Condition 4 implies 1. Set $X = 2$. ∎

Note: The above theorem is adapted from [20].

OBSERVATION 4.5. Inconsistent Theories. The previous result makes it easy to identify the inconsistent theories. The existence of e_X implies that TX is non-empty if X is. If **T** is inconsistent, it follows that TX has exactly one element if X is non-empty. Thus there are at most two ways to define T according as $T\varnothing$ is empty or has one element. In either case, e and $(-)^{\#}$ are uniquely defined and satisfy the three axioms. Every inconsistent theory is isomorphic (in the sense of 1.19) to one of these two. These theories are uninteresting and we shall largely forget about them, adapting our notations to the consistent case. In particular, for $x \in X$ we shall write $x \in TX$ instead of the more cumbersome $e_X(x)$ in most cases.

DEFINITION 4.6. Noise-Free Theories. The set $T\varnothing$ of "distributions on no outcomes" represents "noise." **T** is *noise-free* if $T\varnothing = \varnothing$.

DEFINITION 4.7. Theories with Crisp Points. A "point" is a distribution on one outcome. Let 1 be a one-element set. **T** *has crisp points* if $T1 = 1$ (more precisely: $e_1 : 1 \to T1$ is bijective). By 4.3 it does not matter which one-element set we choose.

THEOREM 4.8. *A consistent theory with crisp points is noise-free.*

Proof. Consider the square shown below in which φ is the unique homomorphism from $T\varnothing$ to $T1$. The square commutes because there is only

$$\begin{tikzcd} T\phi \arrow[r, "\varphi"] \arrow[d, "\varphi"'] & T1 \arrow[d, "T(\text{true})"] \\ T1 \arrow[r, "T(\text{false})"'] & T2 \end{tikzcd}$$

one homomorphism from $T\varnothing$ to $T2$. Since **T** has crisp points $T(\textbf{true}) = \textbf{true}$ and $T(\textbf{false}) = \textbf{false}$. Since $\textbf{true} \neq \textbf{false}$, $T\varnothing$ must be empty. ∎

EXAMPLE 4.9. Pure Noise. Let N be any set (of "pure noises"). Define a fuzzy theory **T** by $TX = X + N$ (+ indicates disjoint union), $e_X(x) = x$, $\alpha^{\#}(x) = \alpha(x)$ if $x \in X$ and $= x$ if $x \in N$. Thus a distribution is either crisp or a pure noise. $T\varnothing = N$. This theory is anti-reflexive and faithful and satisfies the eigenstate condition. The equality function is given by

$$
\begin{aligned}
eq_X(p,q) &= \mathbf{true} && \text{if } p, q \in X,\ p = q \\
&= \mathbf{false} && \text{if } p, q \in X,\ p \neq q \\
&= p && \text{if } p \in N \\
&= q && \text{if } p \in X,\ q \in N.
\end{aligned}
$$

Thus **T**-equality is not symmetric unless N has only one element.

The next result makes use of the degree-of-membership map dm_X: $2^X \times TX \to T2$ of 2.2.

THEOREM 4.10. **T** *has crisp points if and only if for every* X *and* $p \in TX$, $dm_X(X, p) = \mathbf{true}$.

Proof. If $dm_X(X, p)$ is always **true**, argue as follows. The map **true**: $1 \to 2$ is injective so, by 4.3, $T(\mathbf{true})$: $T1 \to T2$ is injective. Since $\mathbf{true} = \chi_1 : 1 \to 2$, $T(\mathbf{true}) = \mathrm{dm}_1(1, -)$. Thus $T(\mathbf{true})$: $T1 \to T2$ is injective and has a one-element image which implies that $T1$ has one element. Conversely, assume $T1$ has one element. Consider the triangles below. Here f is the unique

$$
\begin{tikzcd}
X \arrow[r, "f"] \arrow[dr, "\chi_X"'] & 1 \arrow[d, "\mathbf{true}"] \\
& 2
\end{tikzcd}
\qquad\qquad
\begin{tikzcd}
TX \arrow[r, "Tf"] \arrow[dr, "T\chi_X"'] & T1 = 1 \arrow[d, "T(\mathbf{true})"] \\
& T2
\end{tikzcd}
$$

function from X to 1. The leftmost triangle commutes because χ_X is constantly **true**. The rightmost triangle results by applying functoriality. Since $T1 = 1$, $dm_X(X, -) = T\chi_X$ has image $\{\mathbf{true}\}$ as desired. ∎

The following is adapted from [27, p. 23].

THEOREM 4.11. *Every fuzzy theory has a largest canonical subtheory with crisp points.*

Proof. If **S** is a canonical subtheory of **T**, it is immediate from the definitions that for any X, $dm_X(X, -)$ for **T** restricted to SX is the same map as $dm_X(X, -)$ for **S**. It then follows from 4.10 that if CX is defined as the subset of all $p \in TX$ with $dm_X(X, p) = \mathbf{true}$, then **S** has crisp points if and only if $SX \subset CX$ for all X. To complete the proof we must prove that C is a subtheory. To see that $e_X(x) \in CX$ use (B) as follows: $dm_X(X, e_X(x)) = (T\chi_X)(e_X(x)) = e_2\chi_X(x) = \mathbf{true}$. Now let α_0: $X \to CY$ and let i_A: $CA \to TA$ denote the inclusion map. Consider the diagram shown below. If α: $X \to TY$

is defined to be $i_Y \alpha_0$, we must show that $\alpha^\#$ maps CX into CY. Now observe

$$\begin{tikzcd}
X \arrow[r, "\alpha_0"] \arrow[d] \arrow[ddr, "\alpha"] \arrow[dd, bend right=70, "e_X"'] & CY \arrow[dd, "i_Y"] \\
CX \arrow[ur, dashed] \arrow[d, "i_X"'] & \\
TX \arrow[r, "\alpha^\#"] \arrow[dr, "T\chi_X"'] & TY \arrow[d, "T\chi_Y"] \\
& T2
\end{tikzcd}$$

that $((T\chi_Y)\alpha^\#)e_X = (T\chi_Y)(\alpha^\# e_X) = (T\chi_Y) i_Y \alpha_0$ is constantly **true** because $(T\chi_Y) i_Y$ is, whereas $(T\chi_X)e_X$ is also constantly **true** as was shown a few lines above. Applying (D) and 4.1.2, $(T\chi_Y)\alpha^\# = T\chi_X$. Thus for $p \in TX$, $dm_Y(Y, \alpha^\#(p)) = (T\chi_Y)(\alpha^\#(p)) = (T\chi_X)(p) = dm_X(X, p)$. In particular, if $p \in CX$, $\alpha^\#(p) \in CY$. ∎

EXAMPLES 4.12. Examples of the Tf construction are shown in the table below.

Theory	$(Tf)(p)$ for $f: X \to Y$, $p \in TX$
Crisp set theory	$f(p)$
Fuzzy set theory	$y \mapsto \mathrm{Sup}(p(x): f(x) = y)$
Probabilistic set theory	$y \mapsto \sum (p(x): f(x) = y)$
Possibilistic set theory	$\{f(x): x \in p\}$
Credibility theory	$(c, f(x))$ if $p = (c, x)$
Priority theory	"Leftmost occurrence reduction" of $f(x_1) \cdots f(x_n)$ if $p = x_1 \cdots x_n$
Neighborhood theory	$\{B \subset Y: \{x \in X: f(x) \in B\} \in p\}$

Crisp set theory, probabilistic set theory, priority theory and neighborhood theory have crisp points. The subset of TX comprising the largest subtheory with crisp points in the remaining three examples is as follows. For fuzzy sey theory it is all p with $\mathrm{Sup}_x p(x) = 1$. This coincides with the "normalized fuzzy variables" mentioned in 1.15 when X is finite, but is generally a larger subtheory. For possibilistic set theory it is all non-empty subsets. For credibility theory it is the distributions of form $(1, x)$ so that in this case the largest subtheory with crisp points is isomorphic to crisp set theory.

5. Independent Joint Distributions

In this section we study multivariable mappings which are homomorphic in each variable separately and use them to characterize commutative theories. Commutative theories admit a concept of independence for joint distributions and are characterized by the commutativity of each pair of distributions.

Definition 5.1. For $n \geqslant 1$, a function $\varphi\colon TX_1 \times \cdots \times TX_n \to TY$ is an *n-homomorphism* if for each $i \in \{1, \ldots, n\}$ and for each $p_j \in TX_j$ (all $j \neq i$), the function $\psi\colon TX_i \to TY$, $\psi(q) = \varphi(p_1, \ldots, p_{i-1}, q, p_{i+1}, \ldots, p_n)$ is a homomorphism. Thus a 1-homomorphism is a homomorphism.

Notation 5.2. If at all possible we shall use simply $e\colon X_1 \times \cdots \times X_n \to TX_1 \times \cdots \times TX_n$ for the more cumbersome $e_{X_1} \times \cdots \times e_{X_n}$.

Theorem 5.3. *Given two n-homomorphisms $\varphi_1, \varphi_2\colon TX_1 \times \cdots \times TX_n \to TY$ such that $\varphi_1 e = \varphi_2 e\colon X_1 \times \cdots \times X_n \to TY$, $\varphi_1 = \varphi_2$.*

Proof. Use induction on n. For $n = 1$ this is just 4.1.2. Now suppose that there exist $p_i \in TX_i$ with $\varphi_1(p_1, \ldots, p_{n+1}) \neq \varphi_2(p_1, \ldots, p_{n+1})$. For $j = 1, 2$, set $\psi_j(p) = \varphi_j(p_1, \ldots, p_n, p)$. As $\psi_1 \neq \psi_2$ and both are homomorphisms, there exists $x_{n+1} \in X_{n+1}$ with $\psi_1(x_{n+1}) \neq \psi_2(x_{n+1})$. For $j = 1, 2$ define $\gamma_j\colon TX_1 \times \cdots \times TX_n \to TY$ by $\gamma_j(q_1, \ldots, q_n) = \varphi_j(q_1, \ldots, q_n, x_{n+1})$. Then the γ_j are n-homomorphisms and, setting $q_i = p_i$, $\gamma_1 \neq \gamma_2$. By the induction hypothesis, there exists $x_i \in X_i$ for $i = 1, \ldots, n$ such that $\gamma_1(x_1, \ldots, x_n) \neq \gamma_2(x_1, \ldots, x_n)$ as desired. ∎

Observation 5.4. Two Candidates. As discussed in the Introduction, $T(X \times Y)$ is the set of "joint distributions on X, Y" and, in isolating the concept of "independence" for such joint distributions, it is natural to seek a map of form $TX \times TY \to T(X \times Y)$ whose image constitutes the independent ones. We observe here that there are in fact two candidates Γ_1, Γ_2 for such a map. The construction uses the exponential laws 2.3 and coordinatewise extension 2.4.

$X \times Y \xrightarrow{e_{X \times Y}} T(X \times Y)$	Step 1	$X \times Y \xrightarrow{e_{X \times Y}} T(X \times Y)$
$X \xrightarrow{\Gamma_1^2} T(X \times Y)^Y$	Step 2	$Y \xrightarrow{\Gamma_2^2} T(X \times Y)^X$
$TX \xrightarrow{\Gamma_1^3} T(X \times Y)^Y$	Step 3	$TY \xrightarrow{\Gamma_2^3} T(X \times Y)^X$
$TX \times Y \xrightarrow{\Gamma_1^4} T(X \times Y)$	Step 4	$X \times TY \xrightarrow{\Gamma_2^4} T(X \times Y)$

$$Y \xrightarrow{\Gamma_1^5} T(X \times Y)^{TX} \qquad \text{Step 5} \qquad X \xrightarrow{\Gamma_2^5} T(X \times Y)^{TY}$$

$$TY \xrightarrow{\Gamma_1^6} T(X \times Y)^{TX} \qquad \text{Step 6} \qquad TX \xrightarrow{\Gamma_2^6} T(X \times Y)^{TY}$$

$$TX \times TY \xrightarrow{\Gamma_1} T(X \times Y) \qquad \text{Step 7} \qquad TX \times TY \xrightarrow{\Gamma_2} T(X \times Y)$$

EXAMPLE 5.5. Priority Theory. For the theory of 1.8,

$$\Gamma_1(x_1 \cdots x_m, y_1 \cdots y_n) = (x_1, y_1) \cdots (x_1, y_n) \cdots (x_m, y_1) \cdots (x_m, y_n),$$

$$\Gamma_2(x_1 \cdots x_m, y_1 \cdots y_n) = (x_1, y_1) \cdots (x_m, y_1) \cdots (x_1, y_n) \cdots (x_m, y_n).$$

Hence $\Gamma_1 \neq \Gamma_2$.

EXAMPLE 5.6. Commuting Distributions. For any sets X, m, n there are canonical isomorphisms $(X^m)^n \cong X^{m \times n} \cong (X^n)^m$ using the exponential laws 2.3 wherein $(x_{ij}: i \in m, j \in n) \in X^{m \times n}$ corresponds to $((X_{ij}: i \in m): i \in n) \in (X^m)^n$ and to $((x_{ij}: j \in n): i \in m) \in (X^n)^m$. Say that $p \in Tm$, $q \in Tn$ *commute* if their corresponding operations of 3.1 do, that is, if for every set X diagram (A) commutes. Equivalently, given $(r_{ij}: i \in m, j \in n) \in (TX)^{m \times n}$,

$$\begin{tikzcd}
(TX)^{m \times n} \arrow[r] \arrow[d] & ((TX)^m)^n \arrow[r, "(\hat{p}_X)^n"] & (TX)^n \arrow[dd, "\hat{q}_X"] \\
((TX)^n)^m \arrow[d, "(\hat{q}_X)^m"'] & & \\
(TX)^m \arrow[rr, "\hat{p}_X"'] & & TX
\end{tikzcd} \qquad \text{(A)}$$

$\hat{q}_X(\hat{p}_X(r_{ij}: i \in m): j \in n) = \hat{p}_X(\hat{q}_X(r_{ij}: j \in n): i \in m)$.

For example, a binary operation $+$ and a ternary operation $*$ commute if $(a_{11} + a_{21}) * (a_{12} + a_{22}) * (a_{13} + a_{23}) = (a_{11} * a_{12} * a_{13}) + (a_{21} * a_{22} * a_{23})$.

DEFINITION AND THEOREM 5.7. The following conditions on a fuzzy theory **T** are equivalent and define when **T** is a *commutative* theory.

1. For all sets $X, Y, e_{X \times Y}: X \times Y \to T(X \times Y)$ has a 2-homomorphic extension $\Gamma: TX \times TY \to T(X \times Y)$. Such Γ (unique by 5.3) is the *independent joint distributions map.*

2. Every function of form $\alpha: X_1 \times \cdots \times X_n \to TY$ $(n \geqslant 1)$ has a unique n-homomorphic extension $\bar{\alpha}: TX_1 \times \cdots \times TX_n \to TY$.

3. For all sets X, Y the two maps Γ_1, Γ_2 of 5.4 are equal.

4. Every pair of **T**-distributions commutes.

Proof. Condition 1 is equivalent to 2. Condition 1 is a special case of 2. Conversely, use induction on n. For $n=1$ use $\alpha^{\#}$. Given α: $X_1 \times \cdots \times X_n \to TY$, use the inductive hypothesis to obtain an n-homomorphism ψ: $TX_1 \times \cdots \times TX_n \to T(X_1 \times \cdots \times X_n)$ extending $e_{X_1} \times \cdots \times e_{X_n}$. It is easily checked that the composition is an $(n+1)$-

$$\begin{tikzcd}
TX_1 \times \cdots \times TX_n \times TX_{n+1} \arrow[r, "\psi \times id"] \arrow[ddr] & T(X_1 \times \cdots \times X_n) \times TX_{n+1} \arrow[d, "\Gamma"] \\
 & T(X_1 \times \cdots \times X_{n+1}) \arrow[d, "\alpha^{\#}"] \\
 & TY
\end{tikzcd}$$

homomorphism extending α. Uniqueness follows from 5.3.

Condition 1 is equivalent to 3. If $\Gamma_1 = \Gamma_2$, the result is a 2-homomorphic extension of e since Γ_1 is homomorphic in the second variable and Γ_2 is homomorphic in the first variable (Step 6 of 5.4). Conversely, given Γ, proceed as follows. Consider the seven steps in the definition of Γ_1 but working backwards starting with Γ, proceeding from Step 6 to Step 5 by composing with e_Y and proceeding from Step 3 to Step 2 by composing with e_X. Then Γ_1^6 is a homomorphism coordinatewise since Γ is 2-homomorphic, so that $\Gamma_1^6 = (\Gamma_1^5)^{\#}$. Similarly, Γ_1^3 is a homomorphism coordinatewise since Γ is 2-homomorphic and Γ_1^4 is just a restriction of Γ, so $\Gamma_1^3 = (\Gamma_1^2)^{\#}$. But then it is clear that $\Gamma = \Gamma_1$. Similarly, $\Gamma = \Gamma_2$.

Condition 3 is equivalent to 4. Let $p \in Tm$, $q \in Tn$ and set $s = \Gamma_1(p, q) \in T(m \times n)$. We begin by using 5.4 (but writing m, n instead of X, Y) to compute the operation $\hat{s}_X$: $(TX)^{m \times n} \to TX$ induced by s. Let $(r_{ij}$: $i \in m$, $j \in n) \in (TX)^{m \times n}$. By Definition 3.1, $\hat{s}_X(r_{ij})$ is obtained by evaluating

$$Tm \times Tn \xrightarrow{\Gamma_1} T(m \times n) \xrightarrow{(r_{ij})^{\#}} TX$$

at (p, q). As $\Gamma_1^6 = (\Gamma_1^5)^{\#}$ is coordinatewise a homomorphism, if we hold p fixed in the map above, we get a homomorphism φ: $Tn \to TX$ so that $\varphi = (s_j : j \in n)^{\#}$, where $s_j = (r_{ij})^{\#}\Gamma_1^4(p, j) \in TX$. Letting in_j: $m \to m \times n$ be the

$$\begin{tikzcd}
Tm \arrow[r, "T\, in_j"] & T(m \times n) \arrow[r, "(r_{ij}\,:\, i \in m,\ j \in n)^{\#}"] & TX \\
m \arrow[u, "e_m"] \arrow[r, "in_j"'] \arrow[ur, "\Gamma_1^2(-,j)"] \arrow[rr, bend right, "(r_{ij}\,:\, i \in m)"'] & m \times n \arrow[u, "e_{m \times n}"'] \arrow[ur, "(r_{ij})"'] &
\end{tikzcd} \qquad \text{(B)}$$

injection map $i \mapsto (i, j)$, diagram (B) commutes. Using this diagram and the associativity axiom for $\mathbf{T}$, $s_j = (r_{ij}: i \in m, j \in n)^{\#}(\Gamma_1^2)^{\#}(p)(j) = ((r_{ij}: i \in m, j \in n)^{\#}(\Gamma_1^2)(-, j))^{\#}(p) = (r_{ij}: i \in m)^{\#}(p) = \hat{p}_X(r_{ij}: i \in m)$. We then compute that $\hat{s}_X(r_{ij}: i \in m, j \in n) = (r_{ij}: i \in m, j \in n)^{\#}\Gamma_1(p, q) = \varphi(q) = (s_j: j \in n)^{\#}(q) = \hat{q}_X(s_j: j \in n) = \hat{q}_X(\hat{p}_X(r_{ij}: i \in m): j \in n)$. A similar calculation shows that if $t = \Gamma_2(p, q) \in T(m \times n)$ then $\hat{t}_X(r_{ij}: i \in m, j \in n) = \hat{p}_X(\hat{q}_X(r_{ij}: j \in n): i \in m)$. It then follows at once that 3 implies 4. That 4 implies 3 is proved the same way since if p, q commute then $(\Gamma_1(p, q))\hat{} = (\Gamma_2(p, q))\hat{}$ so that $\Gamma_1(p, q) = \Gamma_2(p, q)$ by Theorem 3.3. ∎

EXAMPLE 5.8. The five examples tabulated below are commutative.

Commutative theory	Γ and $\bar{\alpha}$
Crisp set theory	$\Gamma(p, q) = (p, q)$, $\bar{\alpha} = \alpha$
Fuzzy set theory	$\Gamma(p, q)(x, y) = \mathrm{Min}(p(x), q(y)$. $\quad \bar{\alpha}(p_1, \ldots, p_n)(y) = \sup_{x_i \in X_i} \mathrm{Min}(p_1(x_1), \ldots, p_n(x_n), \alpha(x_1, \ldots, x_n)(y))$
Probabilistic set theory	$\Gamma(p, q)(x, y) = p(x)\, q(y)$. $\quad \bar{\alpha}(p_1, \ldots, p_n)(y) = \sum_{x_i \in X_i} p_1(x_1) \cdots p_n(x_n)\, \alpha(x_1, \ldots, x_n)(y)$
Possibilistic set theory	$\Gamma(p, q) = p \times q$. $\quad \bar{\alpha}(p_1, \ldots, p_n) = \bigcup_{x_i \in p_i} \alpha(x_1, \ldots, x_n)$
Credibility theory	$\Gamma((c_1, x_1), (c_2, x_2)) = (\mathrm{Min}(c_1, c_2), (x_1, x_2))$. If $p_i = (c_i, x_i)$ and $\alpha(x_1, \ldots, x_n) = (c, y)$, then $\bar{\alpha}(p_1, \ldots, p_n) = (\mathrm{Min}(c, c_1, \ldots, c_n), y)$

THEOREM 5.9. *If $\mathbf{T}$ is a commutative theory, $\mathbf{T}$-equality is symmetric.*

Proof. By an argument similar to the proof of equivalence of 1 and 3 in 5.7, the $\mathbf{T}$-equality map of 2.5 eq_X: $TX \times TX \to \mathscr{T}$ is the unique 2-homomorphic extension of δ: $X \times X \to \mathscr{T}$, where $\delta(x, y) = \mathbf{true}$ if $x = y$, $= \mathbf{false}$ if $x \neq y$. Since $h(p, q) = eq_X(q, p)$ is another 2-homomorphic extension of δ, $h = eq_X$. ∎

EXAMPLE 5.10. The neighborhood theory and the pure noise theory are not commutative since we have seen in 2.15 and 4.9 that equality is not symmetric in these theories. The converse of 5.9 fails since the priority theory is not commutative by 5.5 whereas its equality map, discussed in 2.11, is symmetric.

Given probability distributions p, q each of p, q may be recovered from their induced independent joint distribution; for example, $p(x) = \sum_y p(x)\, q(y)$. This is not true for fuzzy sets since if $r = \Gamma(p, q)$ with $p(x) \leqslant a$ and $q(y) \geqslant a$ then $r(x, y) = \mathrm{Min}(p(x), q(y)) = p(x)$ is independent of q. No such example exists when $\mathrm{Sup}(p(x)) = 1 = \mathrm{Sup}(q(y))$ in view of the following result.

THEOREM 5.11. *Let* $\mathbf{T}$ *be a commutative theory. Then the following conditions are equivalent.*

1. $\mathbf{T}$ *has crisp points.*

2. *For each pair of sets* X, Y *the independent joint distribution map* Γ: $TX \times TY \to T(X \times Y)$ *is injective.*

Proof. Condition 2 implies 1. Let $1 = \{a\}$ be a one-element set. Let f: $1 \times 1 \to 1$ be the unique map, $f(a, a) = a$. Define

$$\psi = T1 \times T1 \xrightarrow{\Gamma} T(1 \times 1) \xrightarrow{Tf} T1.$$

Since Tf is an isomorphism by 4.3, ψ is injective. As $\mathbf{T}$ is commutative, $\psi(a, -)$: $T1 \to T1$ is a homomorphism. As id_{T1} is the only homomorphism $T1 \to T1$ mapping a to a, $\psi(a, p) = p$ for all $p \in T1$. Similarly, $\psi(p, a) = p$ for all p. But then given $p \in T1$, $\psi(p, a) = \psi(a, p)$ implies $(p, a) = (a, p)$ and hence that $p = a$, that is, $T1 = 1$.

Condition 1 implies 2. We begin by observing that for any theory with crisp points, the homomorphic extension of a constant map is constant. To prove this, let γ: $1 \to TY$ take the value $r \in TY$ and let β: $X \to 1$ be the unique map so that $\alpha = \gamma\beta$: $X \to TY$ is constantly r. As $\gamma^{\#} T\beta$ is a homomorphism

$$\begin{tikzcd}
TX \arrow[r, "T\beta"] & T1 \arrow[r, "\gamma^{\#}"] & TY \\
X \arrow[u, "e_X"] \arrow[r, "\beta"'] & 1 \arrow[u, "e_1"'] \arrow[ur, "\gamma"'] &
\end{tikzcd}$$

equalling $\gamma\beta$ when preceded by e_X, $\gamma^{\#} T\beta = \alpha^{\#}$. But since e_1 is bijective, $\gamma^{\#} T\beta$ is constantly r.

To apply this principle, define pr_X: $X \times Y \to X$ by $pr_X(x, y) = x$ and consult diagram (C). Here β is defined as $e_X pr_X$ and α is obtained from β by coordinatewise extension and the exponential laws (cf. Steps 1 to 4 of 5.4). Since $\beta(x, y) = e_X(x)$ is independent of y and since the homomorphic extension of a constant is constant, $\alpha(x, q) = e_X(x)$ for all $q \in TY$. But then

$$\begin{tikzcd}
TX \times TY \arrow[r, "\Gamma"] & T(X \times Y) \arrow[r, "T pr_X"] & TX \\
X \times TY \arrow[u, "e_X \times id_{TY}"] \arrow[urr, "\alpha"] & (*) & \\
X \times Y \arrow[u, "id_X \times e_Y"] \arrow[uu, bend left=60, "e"'] \arrow[uur, "e_{X \times Y}"] \arrow[uurr, "\beta"'] \arrow[rr, "pr_X"'] & & X \arrow[uu, "e_X"']
\end{tikzcd} \qquad (C)$$

as the perimeter of (C) commutes and as Γ is 2-homomorphic, $(T pr_X)\Gamma$ is the 2-homomorphic extension of β. By uniqueness of homomorphic extension, it follows that triangle $(*)$ commutes. Putting these facts together, $(T pr_X)\,\Gamma(p, q) = e_X^{\#}(p) = p$ for all q. Arguing similarly, $(T pr_Y)\,\Gamma(p, q) = q$. Thus (p, q) may be recovered from $\Gamma(p, q)$ and Γ is injective. ∎

6. The Logic of Propositions

For commutative theories, algebraic operations extend from X to TX and, in particular, every Boolean polynomial extends to $\mathscr{T}$. While $\mathscr{T}$ need not be a Boolean algebra, a large class of Boolean equations continue to hold.

Observation 6.1. The Fuzzification Principle. If $\mathbf{T}$ is a commutative theory, each function f: $X_1 \times \cdots \times X_n \to Y$ has an n-homomorphic extension $\tilde{f}$: $TX_1 \times \cdots \times TX_n \to TY$ defined in the notation of 5.7.2 by $\tilde{f} = \overline{e_Y f}$. When $n = 1$, $\tilde{f} = Tf$. The case $n = 0$ is not covered. Here, f amounts to an element y of Y and we shall define $\tilde{y} = y$ (that is, $e_Y(y)$).

Definition 6.2. The *Boolean logic* of the commutative theory $\mathbf{T}$ is its set $\mathscr{T}$ of $\mathbf{T}$-truth values together with the operations $\tilde{f}$ as f ranges over the finitary Boolean polynomials $2^n \to 2$. The usual practice of defining a Boolean algebra in terms of a small finite set of operations depends on equations of form $f = e$, where f is a Boolean polynomial and e is an expression built from the given small set of operations. Since $\tilde{f} = \tilde{e}$ may not hold in $\mathscr{T}$, it would be prejudicial to favor some operations over others.

In our examples we will emphasize the familiar "or," "and" and "not" operations, written $\vee, \wedge$: $2 \times 2 \to 2$, $(-)'$: $2 \to 2$ as well as the *Boolean conditional* bc_X: $2 \times X \times X \to X$ defined by $bc_X(\textbf{true}, x, y) = x$, $bc_X(\textbf{false}, x, y) = y$. (It is natural to say $bc_X(p, x, y) =$ "if p then x else y"; the more precise notation is necessary owing to a competing conditional to be introduced in Section 8.) Then $\widetilde{bc}_X$: $\mathscr{T} \times TX \times TX \to TX$ is always defined and $\widetilde{bc}_{\mathscr{T}}$ is a ternary operation on $\mathscr{T}$.

EXAMPLE 6.3. Boolean Logic for Fuzzy Sets. See 2.7 for notation.

$$\tilde{\mathbf{true}}_{\mathbf{true}} = 1, \quad \tilde{\mathbf{true}}_{\mathbf{false}} = 0; \quad \tilde{\mathbf{false}}_{\mathbf{true}} = 0, \quad \tilde{\mathbf{false}}_{\mathbf{false}} = 1.$$

In general, for any $f\colon 2^n \to 2$, $t_1,\dots, t_n \in \mathscr{E}$, $x_1,\dots, x_n \in 2$,

$$(\tilde{f}(t_1,\dots, t_n))_k = \operatorname*{Sup}_{f(x_1,\dots,x_n)=k} \operatorname{Min}(t_1(x_1),\dots, t_n(x_n))$$

for $k = \mathbf{true}$, $\mathbf{false}$, and $t_i(x_i)$ alternate notation for subscript notation. For s, $t \in \mathscr{E}$ we will write $s \vee t$ rather than the more cumbersome $s \mathbin{\tilde{\vee}} t$ both here and below, and similarly for $\wedge$ and $'$. Then

$$\begin{aligned}
(s \vee t)_{\mathbf{true}} &= \operatorname{Max}(\operatorname{Min}(s_{\mathbf{true}}, t_{\mathbf{true}}), \operatorname{Min}(s_{\mathbf{true}}, t_{\mathbf{false}}), \operatorname{Min}(s_{\mathbf{false}}, t_{\mathbf{true}})),\\
(s \vee t)_{\mathbf{false}} &= \operatorname{Min}(s_{\mathbf{false}}, t_{\mathbf{false}}),\\
(s \wedge t)_{\mathbf{true}} &= \operatorname{Min}(s_{\mathbf{true}}, t_{\mathbf{true}}),\\
(s \wedge t)_{\mathbf{false}} &= \operatorname{Max}(\operatorname{Min}(s_{\mathbf{false}}, t_{\mathbf{false}}), \operatorname{Min}(s_{\mathbf{false}}, t_{\mathbf{true}}), \operatorname{Min}(s_{\mathbf{true}}, t_{\mathbf{false}})),\\
s'_{\mathbf{true}} &= s_{\mathbf{false}}, \qquad s'_{\mathbf{false}} = s_{\mathbf{true}}.
\end{aligned}$$

Further, for $t \in \mathscr{E}$, q, $r \in TX$, the Boolean conditional is given by

$$\widetilde{bc}_X(t, q, r)(x) = \operatorname{Max}(\operatorname{Min}(t_{\mathbf{true}}, q(x), \operatorname*{Sup}_y r(y)), \operatorname{Min}(t_{\mathbf{false}}, \operatorname*{Sup}_y q(y), r(x)).$$

EXAMPLE 6.4. Boolean Logic for Probabilistic Set Theory. See 2.8 for notation. Here $\mathbf{true} = 1$, $\mathbf{false} = 0$. We have

$$\begin{aligned}
s \vee t &= st + s(1-t) + (1-s)t = s + t - st,\\
s \wedge t &= st,\\
s' &= 1 - s,\\
\widetilde{bc}_X(t, q, r) &= tq + (1-t)r.
\end{aligned}$$

EXAMPLE 6.5. Boolean Logic for Possibilistic Set Theory. In the notation of 2.9, $\tilde{\mathbf{true}} = \mathbf{yes}$ and $\tilde{\mathbf{false}} = \mathbf{no}$. $\vee$ and $\wedge$ are commutative (this is always true as will be proved below) and hence are defined by

$$\mathbf{undefined} \vee t = \mathbf{undefined}, \qquad \mathbf{undefined} \wedge t = \mathbf{undefined},$$

$$\mathbf{maybe} \vee \mathbf{maybe} = \mathbf{maybe}, \qquad \mathbf{maybe} \wedge \mathbf{maybe} = \mathbf{maybe}.$$

And, for $t \neq \mathbf{undefined}$,

$$\mathbf{yes} \vee t = \mathbf{yes}, \qquad \mathbf{yes} \wedge t = t,$$

$$\mathbf{no} \vee t = t, \qquad \mathbf{no} \wedge t = \mathbf{no}.$$

Also,

$$\textbf{undefined}' = \textbf{undefined}, \qquad \textbf{maybe}' = \textbf{maybe},$$

$$\textbf{yes}' = \textbf{no}, \qquad \textbf{no}' = \textbf{yes}.$$

The Boolean conditional is given by

$$\widetilde{bc}_X(\textbf{undefined}, q, r) = \varnothing, \qquad \widetilde{bc}_X(\textbf{maybe}, q, r) = q \cup r,$$

$$\widetilde{bc}_X(\textbf{yes}, q, r) = q, \qquad \widetilde{bc}_X(\textbf{no}, q, r) = r$$

if both q, $r \neq \varnothing$ whereas $\widetilde{bc}_X(t, q, r) = \varnothing$ if either $q = \varnothing$ or $r = \varnothing$.

EXAMPLE 6.6. Boolean Logic for Credibility Theory. $\textbf{true}^{\sim} = (1, \textbf{true})$, $\textbf{false}^{\sim} = (1, \textbf{false})$ where 1 is the greatest element of C). In general, if $f(x_1,..., x_n) = y$, $\tilde{f}((c_1, x_1),..., (c_n, x_n)) = (\text{Min}(c_1,..., c_n), y)$. For $u, v \in 2$, c, $d \in C$ we have

$$(c, u) \vee (d, v) = (\text{Min}(c, d), u \vee v),$$

$$(c, u) \wedge (d, v) = (\text{Min}(c, d), u \wedge v),$$

$$(c, u)' = (c, u').$$

The Boolean conditional is given by

$$\widetilde{bc}_X((c, \textbf{true}), (c_1, x_1), (c_2, x_2)) = (\text{Min}(c, c_1), x_1),$$

$$\widetilde{bc}_X((c, \textbf{false}), (c_1, x_1), (c_2, x_2)) = (\text{Min}(c, c_2), x_2).$$

We wish to motivate what comes next by considering an arbitrary binary operation $f: X^2 \to X$ and its extension $\tilde{f}: (TX)^2 \to TX$, where $\mathbf{T}$ is possibilistic set theory. Then by 5.8, $\tilde{f}(A, B) = \{f(a, b): a \in A, b \in B\}$. Define φ: $(TX)^3 \to TX$ by $\varphi(A, B, C) = \tilde{f}(A, \tilde{f}(B, C)) = \{f(a, f(b, c)): a \in A, b \in B, c \in C\}$. It is clear that φ is a 3-homomorphism (i.e., by 3.7, $\varphi(\bigcup A_i, B, C) = \bigcup \varphi(A_i, B, C)$, etc.). It follows similarly that $\tilde{f}(\tilde{f}(A, B), C)$ is a 3-homomorphism. It is then immediate from Theorem 5.3 that $\tilde{f}$ is associative if f is.

The argument breaks down if a variable is repeated. For example, consider ψ: $(TX)^2 \to TX$, $\psi(A, B) = \tilde{f}(A, \tilde{f}(B, A))$. If $a_i \in A_i$ with $a_1 \notin A_2$, $a_2 \notin A_1$ and if $b \in B$, $f(a_1, f(b, a_2)) \in \psi(A_1 \cup A_2, B)$ but will not, in general, be an element of $\psi(A_1, B) \cup \psi(A_2, B)$, so ψ need not be a 2-homomorphism. In such a case, if $g: X^2 \to X$, $g(x, y) = f(x, f(y, x))$, $\psi \neq \tilde{g}$.

DEFINITION 6.7. Universal algebra [7, 15] deals with equationally-definable classes of algebras. The treatment here will be as informal as possible. Examples of algebras defined by operations and equations include groups, rings and Boolean algebras (but not fields since multiplicative inverse is not a totally-defined unary operation). Boolean algebras may be presented by imposing two binary operations $\wedge$, $\vee$ (infimum and supremum), one unary operation $(-)'$ (complement) and two nullary operations (= constants) 0, 1 (the least and greatest elements) and by imposing well-known appropriate equations. In addition to the equations provided by the presentation, many other equations will hold. A number of valid equations appear in (A) and (B) below. The expressions on either side of an equation

$$\begin{aligned} x \wedge y &= y \wedge x, \\ x \wedge (y \wedge z) &= (x \wedge y) \wedge z, \\ x \wedge 1 &= x, \\ x'' &= x, \\ (x \vee y)' &= x' \wedge y', \end{aligned} \tag{A}$$

$$\begin{aligned} x \wedge x &= x, \\ x \vee x &= x, \\ x \vee 1 &= 1, \\ x \vee x' &= 1, \\ x \wedge x' &= 0, \\ x \wedge x' &= y \wedge y', \\ x \wedge (y \vee z) &= (x \wedge y) \vee (x \wedge z) \end{aligned} \tag{B}$$

are called *terms*. Given an equational presentation Ω an *Ω-algebra* is a set equipped with the corresponding operations which satisfy the given equations. If X is an Ω-algebra and if t is a term with n variables, t induces a map $X^n \to X$ by substituting elements of X for the syntactic variables. If X is an Ω-algebra and if $\mathbf{T}$ is a commutative theory then each of the Ω-operations on X extends to TX by the fuzzification principle 6.1 and so each Ω-term with n variables induces a map $(TX)^n \to TX$. Define an Ω-term to be *multi-homomorphic* if it has $n \geqslant 1$ variables and is such that whenever X is an Ω-algebra and $\mathbf{T}$ is a commutative theory, the induced map $(TX)^n \to TX$ is n-homomorphic. All terms on either side in (A) are multi-homomorphic because of

THEOREM 6.8. *Let t be an Ω-term with $n \geqslant 1$ variables. Then if each variable in t occurs without repetition, t is multi-homomorphic.*

Proof. We use induction on the derivation tree of t. (For example, if $t = x \vee (y \wedge 1)$, the derivation tree of t is

$$\begin{array}{ccccc} & & \vee & & \\ & / & & \backslash & \\ x & & & & \wedge \\ & & & / & & \backslash \\ & & & y & & & 1 \end{array}$$

Derivation trees are unique and, unless they consist of a single variable, have a topmost decoupling; see [7, III.2; 24, 1.11].

Let X be an Ω-algebra and let $\mathbf{T}$ be a commutative theory. Each n-ary operation ω in the presentation Ω induces an operation $\omega_X\colon X^n \to X$ and hence the operation $\tilde{\omega}_X\colon (TX)^n \to TX$. Thus each Ω-term t with n variables has interpretations $t_X\colon X^n \to X$ and $t_{TX}\colon (TX)^n \to TX$. (For example, let Ω present Boolean algebras as in 6.7, let B be a Boolean algebra, let $\mathbf{T}$ be possibility theory; then $\Lambda_B(b, c) = b \wedge c$, $\tilde{\Lambda}_B(L, M) = \{b \wedge c\colon b \in L,\ c \in M\}$; if $t = x \wedge x$ then $t_{TX}(L) = \tilde{\Lambda}_B(L, L) = \{b \wedge c\colon b, c \in L\}$; notice that $\tilde{t}_X(L) = \{b \wedge b\colon b \in L\} = L$ is a different operation.) The statement to be proved by induction is: "if t has $n \geqslant 1$ variables which occur without repetition then $t_{TX} = \tilde{t}_X$."

For the basis step, if the derivation tree of t consists of a single abstract variable, its interpretation in every algebra is the identity map so that $\tilde{t}_X = T(t_X) = T(id_X) = id_{TX} = t_{TX}$.

For the inductive step, if t has $n \geqslant 1$ variables and is not a single variable it has a unique topmost decoupling

$$t = \begin{array}{ccccc} & & \omega & & \\ & / & & \backslash & \\ t1 & & \cdots & & tm \end{array}$$

with $m \geqslant 1$, ω an m-ary operation in the presentation Ω and $t1, \ldots, tm$ terms. Clearly, any repetition of variables in a ti would induce one in t so that the inductive hypothesis applies and $\widetilde{ti}_X = ti_{TX}$ for $1 \leqslant i \leqslant m$. Then $t_{TX} = \tilde{\omega}_X(t1_{TX}, \ldots, tm_{TX}) = \tilde{\omega}_X(\widetilde{t1}_X, \ldots, \widetilde{tm}_X)$. Fix $k \in \{1, \ldots, n\}$ and substitute fixed elements of TX for all but the kth abstract variable in t. Substitute variable $p \in TX$ for the kth abstract variable in t which resides in ti for unique i because there is no repetition of variables in t. Then t_{TX} is a function of p. Since $\widetilde{tj}_X$ is independent of p for $j \neq k$, $t_{TX}(p)$ is the composition of the two homomorphisms $\widetilde{tk}_X(\ldots, p, \ldots)$ and $\omega_X(\widetilde{t1}_X, \ldots, \widetilde{t(k-1)}_X, (-), \widetilde{t(k+1)}_X, \ldots, \widetilde{tm}_X)$. Thus t_{TX} is n-homomorphic. It follows from 5.3 that $t_{TX} = \tilde{t}_X$. ∎

DEFINITION 6.9. A Condition of Eilenberg. Given an equational presentation Ω, an Ω-equation $t = s$ is *nonrepetitive* if the set of abstract variables occurring in t coincides with the set of abstract variables occurring in s and if no repetition of variables occurs either in t or in s. In the examples of 6.7, the equations of (A) are nonrepetitive whereas none of those of (B) is nonrepetitive.

I believe that the linear theories studied by Eilenberg (mentioned in the Introduction) amount to equational presentations in which each equation is nonrepetitive and that Eilenberg proved the following result for the case $\mathbf{T} =$ possibilistic set theory.

METATHEOREM 6.10. *Let X be an Ω-algebra and let* $\mathbf{T}$ *be a commutative theory. Then every nonrepetitive Ω-equation true for X is true for TX.*

Proof. Let $t = s$ be a nonrepetitive equation with common number of variables n. Adopt the notations of the proof of 6.8. If $n > 1$ then by 6.8, t_{TX}, s_{TX} are n-homomorphisms $(TX)^n \to TX$ and so are equal by 5.3 if $t_X = s_X$. Now consider the case $n = 0$. (Example: $(1 \vee 1) \wedge 1 = 1 \vee 0$.) For each m-ary operation ω of the presentation Ω with $m > 1$, diagram (C)

$$\begin{tikzcd} X^m \arrow[r, "\omega_X"] \arrow[d, "e"'] & X \arrow[d, "e_X"] \\ (TX)^m \arrow[r, "\tilde{\omega}_X"'] & TX \end{tikzcd} \tag{C}$$

commutes. It follows by induction on the derivation trees, that $t_{TX} = e_X(t_X)$ and s similarly, so that if $t_X = s_X$, $t_{TX} = s_{TX}$. ∎

EXAMPLE AND OPEN QUESTION 6.11. The equation $x \wedge (x \wedge y) = (x \wedge y) \wedge x$ is not nonrepetitive but is true in the Boolean logic of $\mathbf{T}$ because it is a consequence of the nonrepetitive equation $x \wedge y = y \wedge x$. The equation $x \vee x = x \vee (x' \wedge x)$ is false for the Boolean logic of probabilistic set theory but is true for that of possibilistic set theory. In particular, this equation is not a consequence of nonrepetitive equations. We leave unanswered the basic question: Is every equation true in the Boolean logic of every commutative theory necessarily a consequence of nonrepetitive equations?

DEFINITION 6.12. Ordered Structure. Let $\mathbf{T}$ be a commutative theory. Consider the following three sets of equations for the Boolean logic of $\mathbf{T}$,

$$x \wedge (y \wedge z) = (x \wedge y) \wedge z, \qquad x \wedge y = y \wedge x,$$
$$x \vee (y \vee z) = (x \vee y) \vee z, \qquad x \vee y = y \vee x, \tag{D}$$
$$x'' = x, \qquad (x \wedge y)' = x' \vee y', \qquad (x \vee y)' = x' \wedge y',$$

$$x \wedge x = x, \qquad x \vee x = x, \tag{E}$$

$$x \wedge (x \vee y) = x, \qquad x \vee (x \wedge y) = x. \tag{F}$$

All equations in (D) are nonrepetitive and so hold. By (D) if either equation in (E) holds so does the other, and similarly for (F). If (E) holds then $x \leqslant y$ defined by $x \wedge y = x$ is a partial order with respect to which $x \wedge y = \mathrm{Inf}(x, y)$ and, dually, $x \leqslant' y$ defined by $x \vee y = y$ is a partial order with $x \vee y = \mathrm{Sup}(x, y)$. By (D), $x \leqslant y$ if and only if $y' \leqslant' x'$. If these orders exist, they coincide if and only if (F) holds and in this case we say that the Boolean logic of $\mathbf{T}$ is a *lattice*.

EXAMPLE 6.13.

Commutative theory	Partially-ordered status of Boolean logic
Fuzzy set theory	$s \leqslant t \Leftrightarrow s_{\mathbf{true}} \leqslant t_{\mathbf{true}}$ and ($s_{\mathbf{false}} = \mathrm{Min}(s_{\mathbf{true}}, t_{\mathbf{false}})$ or $\mathrm{Min}(s_{\mathbf{true}}, t_{\mathbf{false}}) \leqslant s_{\mathbf{false}} \leqslant \mathrm{Max}(t_{\mathbf{false}}, t_{\mathbf{true}})$; not a lattice
Probabilistic set theory	Not partially ordered
Possibilistic set theory	**undefined** < **false** < **maybe** < **true**; not a lattice
Credibility theory	$(c, u) \leqslant (d, v)$ if $c \leqslant d$ and ($u = v$ or $u =$ **false**, $v =$ **true**); not a lattice

OPEN QUESTION 6.14. Characterize those commutative theories whose Boolean logic is a lattice with **false** as least element and **true** as greatest element. The following result shows that at the very least such theories have crisp points.

THEOREM 6.15. *For a commutative theory* $\mathbf{T}$, *the following are equivalent.*

1. $\mathbf{T}$ *has crisp points.*
2. *The Boolean logic of* $\mathbf{T}$ *satisfies* $x \wedge \mathbf{false} = \mathbf{false}$.

Proof. Condition 1 implies 2. The map φ: $\mathscr{E} \to \mathscr{E}$ defined by $\varphi(r) = r \wedge \mathbf{false}$ is the unique homomorphism mapping both **true**, **false** to

$$\begin{tikzcd}
\mathscr{E} \arrow[rr, "\varphi"] \arrow[dr, "Tf"'] & & \mathscr{E} \\
& T1 \arrow[ur, "Tg"'] & \\
2 \arrow[uu, "e_2"] \arrow[r, "f"'] & 1 \arrow[u, "e_1"] \arrow[r, "g"'] & 2 \arrow[uu, "e_2"']
\end{tikzcd} \tag{G}$$

false. Define f: $2 \to 1$ to be the unique map to the one-element set $1 = \{a\}$ and define g: $1 \to 2$ by $g(a) = \textbf{false}$. As is clear from diagram (G), φ is a homomorphism satisfying $\varphi e_2 = e_2 gf$ so that $\varphi = T(gf) = TgTf$. Since $T1$ has only one element by hypothesis, φ is constant as desired.

Condition 2 implies 1. Using the same notations, consider diagram (H).

$$\begin{tikzcd}
1 \arrow[r, "g"] \arrow[d, "e_1"'] & 2 \arrow[d, "e_2"] \\
T1 \arrow[r, "Tg"] \arrow[dr, "Tg"'] & \mathscr{E} \arrow[d, "\varphi"] \\
& \mathscr{E}
\end{tikzcd} \tag{H}$$

As $Tg\, e_1(a) = e_2\, g(a) = \textbf{false} = \varphi(\textbf{false}) = \varphi e_2\, g(a) = \varphi Tg\, e_1(a)$, the homomorphisms φTg and Tg are equal. Since φ is constant by hypothesis, so is Tg. By 4.3, Tg is injective. It follows that $T1$ has only one element. ∎

7. Superposition

Matrix theories over a complete partial semiring admit a superposition principle. Fuzzy set theory and possibility theory are examples.

Definition 7.1. Complete Partial Semirings. A *complete partial semiring* is $(R, \Sigma, \cdot, 1)$, where R is a non-empty set, Σ is a partially-defined operation on arbitrary families in R, $\cdot$ is a binary operation on R (which we shall write rs rather than $r \cdot s$) and $1 \in R$ subject to the following four axioms.

Axiom 1. $(rs)t = r(st)$, $r1 = r = 1r$.

Axiom 2. If $\Sigma(r\colon i \in I)$ is defined then for all s, $\Sigma(sr_i\colon i \in I)$, $\Sigma(r_i s\colon i \in I)$ are defined and equal, respectively, $s\sum(r_i\colon i \in I)$, $(\Sigma(r_i\colon i \in I))s$.

Axiom 3. For 1-element families, $\Sigma(r) = r$.

Before stating the next axiom, we point out that by a *partition* on a set I we mean a non-empty family $(I_j\colon j \in J)$ of pairwise disjoint subsets of I whose union is I; but we allow I_j to be empty for any set of j.

Axiom 4. If $(r_i\colon i \in I)$ is any family in R and if $(I_j : j \in J)$ is a partition of I then $\Sigma(r_i\colon i \in I)$ is defined if and only if $(\Sigma(r_i\colon i \in I_j)\colon j \in J)$ is defined and, when defined, they are equal.

This definition is a hybrid of the complete semirings of [8] and the partially-additive semirings of [5]. When Σ is the supremum operation of a complete lattice, we recapture the complete lattice ordered semigroups of [14].

If $\{a\}$ is partitioned into $(\{a\}, \varnothing)$ we deduce for $r \in R$ that $r = \Sigma(r) = \Sigma(\Sigma(r), \Sigma\varnothing)$ so that $\Sigma\varnothing$ exists and acts as an additive zero. We henceforth write 0 for $\Sigma\varnothing$. A similar argument shows that any sum of 0's is 0.

EXAMPLE 7.2. The unit interval is a complete partial semiring in a number of ways. We list four.

1. $\sum = \text{Sup}$, $\cdot = \text{Min}$.
2. $\sum = \text{Sup}$, $\cdot =$ numerical multiplication.
3. For finite families let $\Sigma = \text{Min}(1$, usual sum) and for arbitrary families let $\Sigma(r_i\colon i \in I)$ be the topological limit of the net of finite partial sums. Set $\cdot = \text{Min}$.
4. Σ as in 3 but $\cdot =$ numerical multiplication.

EXAMPLE 7.3. Let $R = \{0, 1\}$. Define $\Sigma(r_i\colon i \in I)$ to be 0, 1, undefined accordingly as $\{i \in I\colon r_i = 1\}$ is empty, has one element, has more than one element. Define $\cdot$ to be the Boolean $\wedge$.

DEFINITION 7.4. Matrix Theories. Let R be a complete partial semiring. The *matrix theory of* R is the fuzzy theory $\mathbf{mat}_R = (T, e, (-)^{\#})$ as follows. $TX = \{p : p$ is a function from X to R such that for every function q from X to R $\Sigma(q(x)\, p(x)\colon x \in X)$ exists in $R\}$. Define $e_X(x)$ to be the Kronecker delta $y \mapsto 1$ if $x = y$, $y \mapsto 0$ if $x \neq y$. Given $\alpha\colon X \to TY$, $p \in TX$, define $\alpha^{\#}(p) = \sum(\alpha_x(y)\, p(x)\colon x \in X)$.

In such a matrix theory, $\alpha\colon X \to TY$ may be thought of as a matrix with entries in R with X indexing columns and Y indexing rows. The composition $\alpha; \beta = \beta^{\#}\alpha$ is then matrix multiplication and e_X is the identity matrix.

It is easy to show that $\mathbf{mat}_R$ is a commutative theory if and only if R is commutative in the sense that $rs = sr$.

EXAMPLE 7.5. The fuzzy set theory $= \mathbf{mat}_R$ for R as in 7.2.1. The possibilistic set theory $= \mathbf{mat}_R$ for R the two-element Boolean lattice with $\Sigma = \vee$, $\cdot = \wedge$. For R as in 7.3, $\mathbf{mat}_R$ is a new theory, the *partial functions theory*. Here, the image of e_X misses only one element of TX, the constantly-0 function which we interpret as "undefined." Thus a matrix α: $X \to TY$ corresponds to the partial function $x \mapsto \alpha(x)$ if $\alpha(x) \in Y$, undefined else; this correspondence is bijective between matrices and partial functions. Matrix multiplication is the usual composition of partial functions. The partial functions theory is a subtheory of possibilistic set theory via the theory map that sends "undefined" to the empty set.

OBSERVATION 7.6. Superposition. Given a function f: $X + Y \to Z$ defined on the disjoint union of X and Y, let f_X denote the restriction of f to X, f_Y similarly. Any matrix theory admits the *superposition maps*

$$\begin{array}{ccc} T(X+Y) & \xrightarrow{s_{XY}} & TX \times TY, \\ f & \longmapsto & (f_X, f_Y). \end{array}$$

The principle here is that a distribution on any set of outcomes induces distributions on each subset of outcomes (measuring the "contribution" of that subset). For each X, Y s_{XY} is injective and bijective precisely when Σ is defined for every pair. Indeed, s_{XY} is bijective for fuzzy set theory and for possibilistic set theory, but not for the partial functions theory.

The maps s_{XY} may be generalized from binary to arbitrary disjoint unions and in this form it is not hard to impose axioms so as to characterize matrix theories. The proof follows closely that of [17, Theorem 2.1 (4 implies 1)] and so will not be given here.

EXAMPLE 7.7. Multiset Theory. Let $R = \{0, 1, 2, \ldots\} \cup \{\infty\}$ with the usual sum and product. The *multiset theory* is the subtheory T of $\mathbf{mat}_R$ with TX the set of functions p: $X \to R$ such that (i) $\{x\colon p(x) \neq 0\}$ is finite and (ii) no $p(x) = \infty$. If $p(x) = n$ the interpretation is "x occurs in p n times."All of $\mathbf{mat}_R$ is a theory with a similar interpretation.

EXAMPLE 7.8. If $\mathbf{T}$ is noise-free and consistent it is not a matrix theory since there is no map $TX \to TX \times T\varnothing$. For this reason, neither crisp set theory nor probabilistic set theory are matrix theories.

8. The Distributional Conditional

Definition 8.1. The Distributional Conditional. Let $\mathbf{T}$ be a (not necessarily commutative) fuzzy theory. The *distributional conditional maps*

$$\mathscr{E} \times TX \times TX \xrightarrow{dc_X} TX$$

are defined by $dc_X(t, q, r) = \hat{t}_X(q, r)$, where $t \in \mathscr{E} = T2$ induces a binary operation $\hat{t}$ as in 3.1. Our convention is "**true**-coordinate first" so that in $s = (q, r)$, $s_{\mathbf{true}} = q$ and $s_{\mathbf{false}} = r$.

Example 8.2.

Theory	$dc_X(t, q, r)$
Crisp set theory	q if $t =$ **true**, r if $t =$ **false**
Fuzzy set theory	s, where $s(x) = \mathrm{Max}(\mathrm{Min}(t_{\mathbf{true}}, q(x)), \mathrm{Min}(t_{\mathbf{false}}, r(x)))$
Probabilistic set theory	$tq + (1 - t)r$
Possibilistic set theory	q if $t =$ **yes**, r if $t =$ **no**, $\varnothing$ if $t =$ **undefined**, $q \cup r$ if $t =$ **maybe**
Credibility theory	(q, c) if $t = (\mathbf{true}, c)$, (r, c) if $t = (\mathbf{false}, c)$
Priority theory	q if $t =$ **true**, r if $t =$ **false**, obtained respectively from qr, rq accordingly as $t =$ **moretruethanfalse** or **morefalsethantrue**, by deleting all repetitions except the leftmost occurrence
Neighborhood theory	q if $t =$ **yes**, r if $t =$ **no**, $q \cap r$ if $t =$ **maybe**
Pure noise theory	q if $t =$ **true**, r if $t =$ **false**, t if $t \in N$

Comparison of dc_X and $\tilde{b}x_X$ for fuzzy set theory and possibility theory makes the following result a likely conjecture.

Theorem 8.3. *If* $\mathbf{T}$ *is a commutative theory, the following are equivalent.*

1. $\mathbf{T}$ *has crisp points.*
2. *The Boolean conditional coincides with the distributional conditional.*

Proof. Condition 1 implies 2. For fixed q, $r \in TX$, $dc_X(-, q, r)$ and $\tilde{b}x_X(-, q, r)$ are both homomorphisms $\mathscr{E} \to TX$ so that by 4.1 it suffices to show these maps agree on **true** and **false**. Since $\mathbf{T}$ has crisp points, every constant map $TX \to TX$ is a homomorphism. This was shown in the proof of 5.11 (1 implies 2). It follows that the first projection $TX \times TX \to TX$ is the 2-homomorphic extension of the first projection $X \times X \to X$, and so must coincide with $\widetilde{bc}_X(\mathbf{true}, -, -)$. Thus $dc_X(\mathbf{true}, q, r) = q = \widetilde{bc}_X(\mathbf{true}, q, r)$. By a similar argument, $dc_X(\mathbf{false}, q, r) = r = \widetilde{bc}_X(\mathbf{false}, q, r)$.

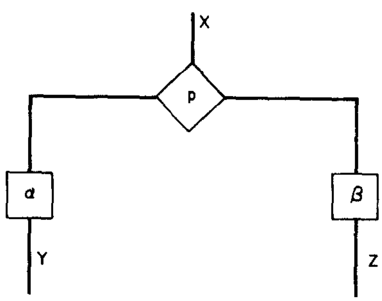
Condition 2 implies 1. By hypothesis, $dc_1 : \mathscr{E} \times T1 \times T1 \to T1$ is a 3-homomorphism so that φ: $T1 \to T1$ defined by $\varphi(r) = dc_1(\mathbf{true}, a, r)$ is a homomorphism. Thus φ is simultaneously the unique homomorphism mapping a to a, namely, id_{T1}, and the map constantly a. If follows that $T1 = 1$. ∎

The examples of 8.2 suggest that the distributional conditional is a little more natural than the Boolean conditional and it is more often defined as well. We favor it in the next two definitions.

DEFINITION 8.4. Given p: $X \to \mathscr{E}$, α: $X \to TY$, β: $X \to TZ$ define

$$X \xrightarrow{\mathbf{if}\, p\, \mathbf{then}\, \alpha\, \mathbf{else}\, \beta} T(Y+Z)$$

by $(\mathbf{if}\ p\ \mathbf{then}\ \alpha\ \mathbf{else}\ \beta)(x) = dc_{Y+Z}(p(x), \alpha(x), \beta(x))$. (Here, if in_Y: $Y \to Y+Z$ is the injection into the disjoint union, in_Z similarly, the more cumbersome but more precise notation is $dc_{Y+Z}(p(x),\ T(in_Y)\,\alpha(x),\ T(in_Z)\,\beta(x))$. Intuition is provided by flowchart (A).



(A)

Define σ^X_{YZ} as the map

$$\mathscr{E}^X \times (TY)^X \times (TZ)^X \xrightarrow{\sigma^X_{YZ}} T(Y+Z),$$
$$(p, \alpha, \beta) \longmapsto \mathbf{if}\ p\ \mathbf{then}\ \alpha\ \mathbf{else}\ \beta$$

and let $\sigma_{YZ} : \mathscr{E} \times TY \times TZ \to T(Y+Z)$ be σ^X_{YZ} when X has one element.

DEFINITION 8.5. **Let T be an arbitrary fuzzy theory. Say that T is *conditional-complete* if σ_{YZ} is surjective whenever both Y, Z are non-empty.** For such **T** it follows that every map $X \to T(Y+Z)$ decomposes into the form **if** p **then** α **else** β.

EXAMPLE 8.6. The theories in the following table are conditional-complete.

Fuzzy theory	$q \in T(Y+Z)$ has form **if** p **then** α **else** β where ...
Crisp set theory	If $q \in Y$, $q =$ **if true then** q **else** β (any $\beta \in Z$). If $q \in Z$, $q =$ **if false then** α **else** q (any $\alpha \in Y$)
Fuzzy set theory	$q =$ **if** $(1, 1)$ **then** q_Y **else** q_Z
Probabilistic set theory	If $k = \sum q_Y = 0$, $q =$ **if** 0 **then** α **else** q_Z(any α). If $k = 1$, $q =$ **if** 1 **then** q_Y **else** β (any β) else $q =$ **if** k **then** $(1/k)q_Y$ **else** $(1/(1-k))q_z$
Possibilistic set theory	$q =$ **if maybe then** $q \cap Y$ **else** $q \cap Z$
Credibility theory	If $q = (y, c)$, $q =$ **if true then** q **else** β (any β). If $q = (z, c)$, $q =$ **if false then** α **else** q (any α).
Pure noise theory	If $q \in Y$, $q =$ **if true then** q **else** β (any β). If $q \in Z$, $q =$ **if false then** α **else** q (any α). If $q \in N$, $q =$ if true then q else q

The formula shown for fuzzy set theory works in fact for any matrix theory. The priority and neighborhood theories are not conditional-complete. If the definitions in 8.4, 8.5 were modified to use the Boolean conditional, fuzzy set theory would not be conditional-complete.

9. CONCLUSIONS

Rather than positing the internal structure of a set of "vague outcomes" in advance, the axioms for a fuzzy theory impose only those structural aspects required to interpret a loop-free program scheme. Just as a discussion of symmetry groups would bypass the use of group theory in arithmetic, our treatment of fuzzy theories bypasses the universal algebra interpretation. Commutative theories with crisp points and commutative matrix theories provide broad classes of examples which are close to standard ones in fuzzy theory. Because a commutative theory allows simultaneous observation of any pair of distributions, any future application to quantum theory is likely to devote attention to noncommutative theories.

REFERENCES

1. S. ALAGIĆ AND M. A. ARBIB, "The Design of Well-Structured and Correct Programs," Springer-Verlag, New York/Berlin, 1978.
2. M. A. ARBIB AND E. G. MANES, Fuzzy machines in a category, *Bull. Austral. Math. Soc.* **13** (1975), 169–210.
3. M. A. ARBIB AND E. G. MANES, A category-theoretic approach to systems in a fuzzy world, *Synthese* **30** (1975), 381–406.
4. M. A. ARBIB AND E. G. MANES, "Arrows, Structures, and Functors: The Categorical Imperative," Academic Press, New York, 1975.
5. M. A. ARBIB AND E. G. MANES, Partially additive categories and flow diagram semantics, *J. Algebra* **62** (1980), 203–227.
6. M. A. ARBIB AND E. G. MANES, The pattern-of-calls expansion is the canonical fixpoint for recursive definitions, *J. Assoc. Comput. Mach.* (April 1982), in press.
7. P. M. COHN, "Universal Algebra," Harper & Row, New York, 1965.
8. S. EILENBERG, "Automata, Languages and Machines," Academic Press, New York, 1974.
9. S. EILENBERG AND J. B. WRIGHT, Automata in general algebras, *Inform. and Control* **11** (1967), 452–470.
10. C. C. ELGOT, Monadic computation and iterative algebraic theories, *in* "Proceedings, Logic Coll. '73" (H. E. Rose and J. C. Shepherdson, Eds.), pp. 175–230, North-Holland, Amsterdam, 1975.
11. B. R. GAINES, Fuzzy reasoning and the logics of uncertainty, *in* "Proceedings, Sixth International Symposium Multiple-Valued Logic," pp. 179–188, IEEE 76CH111-4C, 1976.
12. B. R. GAINES AND L. J. KOHOUT, The fuzzy decade: a bibliography on fuzzy systems and closely related topics, *Internat. J. Man-Mach. Stud.* **9** (1977), 1–68.
13. R. GODEMENT, "Théorie des Faisceaux," Herman, Paris, 1958.
14. J. A. GOGUEN, L-fuzzy sets, *J. Math. Anal. Appl.* **18** (1967), 145–174.
15. G. GRÄTZER, "Universal Algebra," 2nd ed., Springer-Verlag, New York/Berlin, 1979.
16. J. M. JAUCH, "Foundations of Quantum Mechanics," Addison–Wesley, New York, 1968.
17. J. S. JOHNSON AND E. G. MANES, On modules over a semiring, *J. Algebra* **15** (1970), 57–67.
18. P. T. JOHNSTONE, "Topos Theory," Academic Press, New York, 1977.
19. A. KOCK, Bilinearity of cartesian closed monads, *Math. Scand.* **29** (1971), 161–174.
20. F. W. LAWVERE, "Functorial Semantics of Algebraic Theories," Dissertation, Columbia University, 1963.
21. F. E. J. LINTON, Autonomous equational categories, *J. Math. Mech.* **15** (1966), 637–642.
22. S. MAC LANE, "Categories for the Working Mathematician," Springer-Verlag, New York/Berlin, 1971.
23. S. MAC LANE, review of "Topos Theory" by P. T. Johnstone, *Bull. Amer. Math. Soc.* **1** (1979), 1005–1014.
24. E. G. MANES, Algebraic Theories," Springer-Verlag, New York/Berlin, 1976.
25. Z. MANNA, "Mathematical Theory of Computation," McGraw–Hill, New York, 1974.
26. D. S. SCOTT, Continuous lattices, *in* "Toposes, Algebraic Geometry and Logic," Lecture Notes in Mathematics No. 274, pp. 97–136, Springer-Verlag, New York/Berlin, 1972.
27. G. C. WRAITH, "Algebraic Theories," Lectures Notes No. 22, Aarhus Universitet, 1970.
28. L. A. ZADEH, Fuzzy sets, *Inform. and Control* **8** (1965), 338–353.
29. L. A. ZADEH, A fuzzy-algorithmic approach to the definition of complex or imprecise concepts, *Internat. J. Man-Mach. Stud.* **8**(1976), 249–291.